SEQUENCE LISTING

<110> VERTEX PHARMACEUTICALS INC.

<120> METHODS AND COMPOSITIONS FOR RAPID DEVELOPMENT OF
      SCREENING ASSAYS

<130> VPI/02-143WO2

<140> PCT/US03/28078
<141> 2003-09-05

<150> 60/408,297
<151> 2002-09-05

<160> 86

<170> PatentIn version 3.2

<210> 1
<211> 795
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(795)
<223> B-lactamase

<400> 1

```
atg agt cac cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag      48
Met Ser His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln
1               5                   10                  15

ttg ggt gca cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag      96
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys
                20                  25                  30

atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act     144
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr
            35                  40                  45

ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc cgt gtt gac gcc ggg     192
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly
        50                  55                  60

caa gag caa ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt     240
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val
65                  70                  75                  80

gag tac tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta     288
Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val
                85                  90                  95

aga gaa tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc     336
Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala
            100                 105                 110
```

```
aac tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt    384
Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe
        115             120             125

ttg cac aac atg ggg gat cat gta act cgc ctt gat cgt tgg gaa ccg    432
Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro
        130             135             140

gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct    480
Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro
145             150             155             160

gca gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt    528
Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu
                165             170             175

act cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa    576
Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys
                180             185             190

gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att    624
Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile
                195             200             205

gct gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca    672
Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala
        210             215             220

gca ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg    720
Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr
225             230             235             240

acg ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag    768
Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu
                245             250             255

ata ggt gcc tca ctg att aag cat tgg                                 795
Ile Gly Ala Ser Leu Ile Lys His Trp
                260             265
```

<210> 2
<211> 265
<212> PRT
<213> Escherichia coli

<400> 2

```
Met Ser His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln
1               5               10              15

Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys
            20              25              30

Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr
        35              40              45

Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Val Asp Ala Gly
        50              55              60
```

```
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val
65              70              75              80

Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val
            85              90              95

Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala
            100             105             110

Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe
            115             120             125

Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro
    130             135             140

Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro
145             150             155             160

Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu
            165             170             175

Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys
            180             185             190

Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile
            195             200             205

Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala
            210             215             220

Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr
225             230             235             240

Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu
            245             250             255

Ile Gly Ala Ser Leu Ile Lys His Trp
            260             265
```

```
<210> 3
<211> 858
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(858)
<223> B-lactamase

<400> 3
atg aga att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg gca      48
Met Arg Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
1               5               10              15
```

```
ttt tgc ctt cct gtt ttt ggt cac cca gaa acg ctg gtg aaa gta aaa        96
Phe Cys Leu Pro Val Phe Gly His Pro Glu Thr Leu Val Lys Val Lys
            20              25              30

gat gct gaa gat cag ttg ggt gca cga gtg ggt tac atc gaa ctg gat       144
Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
            35              40              45

ctc aac agc ggt aag atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt       192
Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
  · 50              55              60

cca atg atg agc act ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc       240
Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65              70              75              80

cgt gtt gac gcc ggg caa gag caa ctc ggt cgc cgc ata cac tat tct       288
Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
            85              90              95

cag aat gac ttg gtt gag tac tca cca gtc aca gaa aag cat ctt acg       336
Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
            100             105             110

gat ggc atg aca gta aga gaa tta tgc agt gct gcc ata acc atg agt       384
Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
            115             120             125

gat aac act gcg gcc aac tta ctt ctg aca acg atc gga gga ccg aag       432
Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
            130             135             140

gag cta acc gct ttt ttg cac aac atg ggg gat cat gta act cgc ctt       480
Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145             150             155             160

gat cgt tgg gaa ccg gag ctg aat gaa gcc ata cca aac gac gag cgt       528
Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
            165             170             175

gac acc acg atg cct gca gca atg gca aca acg ttg cgc aaa cta tta       576
Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
            180             185             190

act ggc gaa cta ctt act cta gct tcc cgg caa caa tta ata gac tgg       624
Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
            195             200             205

atg gag gcg gat aaa gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg       672
Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
            210             215             220

gct ggc tgg ttt att gct gat aaa tct gga gcc ggt gag cgt ggg tct      720
Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225             230             235             240
```

```
cgc ggt atc att gca gca ctg ggg cca gat ggt aag ccc tcc cgt atc        768
Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
                245              .   250              255

gta gtt atc tac acg acg ggg agt cag gca act atg gat gaa cga aat        816
Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
                260              265              270

aga cag atc gct gag ata ggt gcc tca ctg att aag cat tgg                858
Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
            275              280              285
```

<210> 4
<211> 286
<212> PRT
<213> Escherichia coli

<400> 4

```
Met Arg Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
1               5               10              15

Phe Cys Leu Pro Val Phe Gly His Pro Glu Thr Leu Val Lys Val Lys
            20              25              30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
        35              40              45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
        50              55              60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65              70              75              80

Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
                85              90              95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
            100             105             110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
            115             120             125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
        130             135             140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145             150             155             160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
            165             170             175

Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
            180             185             190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
            195             200             205
```

```
Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
    210             215             220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225             230             235             240

Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
            245             250             255

Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
            260             265             270

Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
        275             280             285
```

```
<210> 5
<211> 795
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(795)
<223> B-lactamase

<400> 5
atg ggg cac cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag      48
Met Gly His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln
1               5               10              15

ttg ggt gca cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag      96
Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys
            20              25              30

atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act     144
Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr
        35              40              45

ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc cgt gat gac gcc ggg     192
Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Asp Asp Ala Gly
    50              55              60

caa gag caa ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt     240
Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val
65              70              75              80

gag tac tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta     288
Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val
            85              90              95

aga gaa tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc     336
Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala
        100             105             110

aac tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt     384
Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe
        115             120             125
```

```
ttg cac aac atg ggg gat cat gta act cgc ctt gat cat tgg gaa ccg        432
Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro
    130             135             140

gag ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct        480
Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro
145             150             155             160

gta gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt        528
Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu
                165             170             175

act cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa        576
Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys
            180             185             190

gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att        624
Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile
        195             200             205

gct gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca        672
Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala
    210             215             220

gca ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg        720
Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr
225             230             235             240

acg ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag        768
Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu
                245             250             255

ata ggt gcc tca ctg att aag cat tgg                                    795
Ile Gly Ala Ser Leu Ile Lys His Trp
            260             265
```

<210> 6
<211> 265
<212> PRT
<213> Escherichia coli

<400> 6

```
Met Gly His Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln
1               5               10              15

Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys
            20              25              30

Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr
            35              40              45

Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Asp Asp Ala Gly
    50              55              60

Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val
65              70              75              80
```

```
Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val
                85                  90                  95

Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala
                100                 105                 110

Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe
            115                 120                 125

Leu His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro
        130                 135                 140

Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro
145                 150                 155                 160

Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu
                165                 170                 175

Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys
            180                 185                 190

Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile
        195                 200                 205

Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala
    210                 215                 220

Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr
225                 230                 235                 240

Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu
            245                 250                 255

Ile Gly Ala Ser Leu Ile Lys His Trp
            260                 265
```

```
<210> 7
<211> 792
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(792)
<223> B-lactamase

<400> 7
atg gac cca gaa acg ctg gtg aaa gta aaa gat gct gaa gat cag ttg      48
Met Asp Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu
1               5                   10                  15

ggt gca cga gtg ggt tac atc gaa ctg gat ctc aac agc ggt aag atc      96
Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile
            20                  25                  30
```

```
ctt gag agt ttt cgc ccc gaa gaa cgt ttt cca atg atg agc act ttt          144
Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe
        35                  40                  45

aaa gtt ctg cta tgt ggc gcg gta tta tcc cgt att gac gcc ggg caa          192
Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln
        50                  55                  60

gag caa ctc ggt cgc cgc ata cac tat tct cag aat gac ttg gtt gag          240
Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu
65                  70                  75                  80

tac tca cca gtc aca gaa aag cat ctt acg gat ggc atg aca gta aga          288
Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg
                85                  90                  95

gaa tta tgc agt gct gcc ata acc atg agt gat aac act gcg gcc aac          336
Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn
                100                 105                 110

tta ctt ctg aca acg atc gga gga ccg aag gag cta acc gct ttt ttg          384
Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu
        115                 120                 125

cac aac atg ggg gat cat gta act cgc ctt gat cat tgg gaa ccg gag          432
His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro Glu
        130                 135                 140

ctg aat gaa gcc ata cca aac gac gag cgt gac acc acg atg cct gta          480
Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val
145                 150                 155                 160

gca atg gca aca acg ttg cgc aaa cta tta act ggc gaa cta ctt act          528
Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr
                165                 170                 175

cta gct tcc cgg caa caa tta ata gac tgg atg gag gcg gat aaa gtt          576
Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val
                180                 185                 190

gca gga cca ctt ctg cgc tcg gcc ctt ccg gct ggc tgg ttt att gct          624
Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala
        195                 200                 205

gat aaa tct gga gcc ggt gag cgt ggg tct cgc ggt atc att gca gca          672
Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala
        210                 215                 220

ctg ggg cca gat ggt aag ccc tcc cgt atc gta gtt atc tac acg acg          720
Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr
225                 230                 235                 240

ggg agt cag gca act atg gat gaa cga aat aga cag atc gct gag ata          768
Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile
                245                 250                 255
```

```
ggt gcc tca ctg att aag cat tgg                                          792
Gly Ala Ser Leu Ile Lys His Trp
            260
```

<210> 8
<211> 264
<212> PRT
<213> Escherichia coli

<400> 8
```
Met Asp Pro Glu Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu
1               5                  10                  15

Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile
            20              25                  30

Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe
            35              40                  45

Lys Val Leu Leu Cys Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln
    50              55                  60

Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu
65                  70                  75                  80

Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg
                85                  90                  95

Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn
            100                 105                 110

Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu
            115                 120                 125

His Asn Met Gly Asp His Val Thr Arg Leu Asp His Trp Glu Pro Glu
    130                 135                 140

Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val
145                 150                 155                 160

Ala Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr
                165                 170                 175

Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val
            180                 185                 190

Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala
            195                 200                 205

Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala
    210                 215                 220

Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr
225                 230                 235                 240

Gly Ser Gln Ala Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile
                245                 250                 255
```

Gly Ala Ser Leu Ile Lys His Trp
            260

<210> 9
<211> 786
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(786)
<223> B-lactamase

<400> 9
atg aaa gat gat ttt gca aaa ctt gag gaa caa ttt gat gca aaa ctc   48
Met Lys Asp Asp Phe Ala Lys Leu Glu Glu Gln Phe Asp Ala Lys Leu
1   5   10   15

ggg atc ttt gca ttg gat aca ggt aca aac cgg acg gta gcg tat cgg   96
Gly Ile Phe Ala Leu Asp Thr Gly Thr Asn Arg Thr Val Ala Tyr Arg
            20   25   30

ccg gat gag cgt ttt gct ttt gct tcg acg att aag gct tta act gta   144
Pro Asp Glu Arg Phe Ala Phe Ala Ser Thr Ile Lys Ala Leu Thr Val
            35   40   45

ggc gtg ctt ttg caa cag aaa tca ata gaa gat ctg aac cag aga ata   192
Gly Val Leu Leu Gln Gln Lys Ser Ile Glu Asp Leu Asn Gln Arg Ile
        50   55   60

aca tat aca cgt gat gat ctt gta aac tac aac ccg att acg gaa aag   240
Thr Tyr Thr Arg Asp Asp Leu Val Asn Tyr Asn Pro Ile Thr Glu Lys
65   70   75   80

cac gtt gat acg gga atg acg ctc aaa gag ctt gcg gat gct tcg ctt   288
His Val Asp Thr Gly Met Thr Leu Lys Glu Leu Ala Asp Ala Ser Leu
                85   90   95

cga tat agt gac aat gcg gca cag aat ctc att ctt aaa caa att ggc   336
Arg Tyr Ser Asp Asn Ala Ala Gln Asn Leu Ile Leu Lys Gln Ile Gly
            100   105   110

gga cct gaa agt ttg aaa aag gaa ctg agg aag att ggt gat gag gtt   384
Gly Pro Glu Ser Leu Lys Lys Glu Leu Arg Lys Ile Gly Asp Glu Val
            115   120   125

aca aat ccc gaa cga ttc gaa cca gag tta aat gaa gtg aat ccg ggt   432
Thr Asn Pro Glu Arg Phe Glu Pro Glu Leu Asn Glu Val Asn Pro Gly
        130   135   140

gaa act cag gat acc agt aca gca aga gca ctt gtc aca agc ctt cga   480
Glu Thr Gln Asp Thr Ser Thr Ala Arg Ala Leu Val Thr Ser Leu Arg
145   150   155   160

```
gcc ttt gct ctt gaa gat aaa ctt cca agt gaa aaa cgc gag ctt tta       528
Ala Phe Ala Leu Glu Asp Lys Leu Pro Ser Glu Lys Arg Glu Leu Leu
                165                 170                 175

atc gat tgg atg aaa cga aat acc act gga gac gcc tta atc cgt gcc       576
Ile Asp Trp Met Lys Arg Asn Thr Thr Gly Asp Ala Leu Ile Arg Ala
                180                 185                 190

gga gcg gca tca tat gga acc cgg aat gac att gcc atc att tgg ccg       624
Gly Ala Ala Ser Tyr Gly Thr Arg Asn Asp Ile Ala Ile Ile Trp Pro
                195                 200                 205

cca aaa gga gat cct gtc ggt gtg ccg gac ggt tgg gaa gtg gct gat       672
Pro Lys Gly Asp Pro Val Gly Val Pro Asp Gly Trp Glu Val Ala Asp
        210                 215                 220

aaa act gtt ctt gca gta tta tcc agc agg gat aaa aag gac gcc aag       720
Lys Thr Val Leu Ala Val Leu Ser Ser Arg Asp Lys Lys Asp Ala Lys
225                 230                 235                 240

tat gat gat aaa ctt att gca gag gca aca aag gtg gta atg aaa gcc       768
Tyr Asp Asp Lys Leu Ile Ala Glu Ala Thr Lys Val Val Met Lys Ala
                245                 250                 255

tta aac atg aac ggc aaa                                               786
Leu Asn Met Asn Gly Lys
                260
```

<210> 10
<211> 262
<212> PRT
<213> Escherichia coli

<400> 10

```
Met Lys Asp Asp Phe Ala Lys Leu Glu Glu Gln Phe Asp Ala Lys Leu
1               5               10                  15

Gly Ile Phe Ala Leu Asp Thr Gly Thr Asn Arg Thr Val Ala Tyr Arg
            20              25                  30

Pro Asp Glu Arg Phe Ala Phe Ala Ser Thr Ile Lys Ala Leu Thr Val
        35              40                  45

Gly Val Leu Leu Gln Gln Lys Ser Ile Glu Asp Leu Asn Gln Arg Ile
    50              55                  60

Thr Tyr Thr Arg Asp Asp Leu Val Asn Tyr Asn Pro Ile Thr Glu Lys
65              70                  75                  80

His Val Asp Thr Gly Met Thr Leu Lys Glu Leu Ala Asp Ala Ser Leu
            85                  90                  95

Arg Tyr Ser Asp Asn Ala Ala Gln Asn Leu Ile Leu Lys Gln Ile Gly
            100                 105                 110

Gly Pro Glu Ser Leu Lys Lys Glu Leu Arg Lys Ile Gly Asp Glu Val
        115                 120                 125
```

```
Thr Asn Pro Glu Arg Phe Glu Pro Glu Leu Asn Glu Val Asn Pro Gly
    130                 135             140

Glu Thr Gln Asp Thr Ser Thr Ala Arg Ala Leu Val Thr Ser Leu Arg
145                 150             155                 160

Ala Phe Ala Leu Glu Asp Lys Leu Pro Ser Glu Lys Arg Glu Leu Leu
            165             170             175

Ile Asp Trp Met Lys Arg Asn Thr Thr Gly Asp Ala Leu Ile Arg Ala
        180             185             190

Gly Ala Ala Ser Tyr Gly Thr Arg Asn Asp Ile Ala Ile Ile Trp Pro
        195             200             205

Pro Lys Gly Asp Pro Val Gly Val Pro Asp Gly Trp Glu Val Ala Asp
    210             215             220

Lys Thr Val Leu Ala Val Leu Ser Ser Arg Asp Lys Lys Asp Ala Lys
225                 230             235                 240

Tyr Asp Asp Lys Leu Ile Ala Glu Ala Thr Lys Val Val Met Lys Ala
            245             250             255

Leu Asn Met Asn Gly Lys
            260
```

<210> 11
<211> 720
<212> DNA
<213> Aequorea victoria

<220>
<221> CDS
<222> (1)..(720)
<223> mutant green fluorescent protein

<400> 11
```
atg gtg agc aag ggc gag gag ctg ttc acc ggg gtg gtg ccc atc ctg    48
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5               10              15

gtc gag ctg gac ggc gac gta aac ggc cac aag ttc agc gtg tcc ggc    96
Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20              25              30

gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg aag ttc atc   144
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35              40              45

tgc acc acc ggc aag ctg ccc gtg ccc tgg ccc acc ctc gtg acc acc   192
Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
        50              55              60
```

```
ttc tcc tac ggc gtg cag tgc ttc agc cgc tac ccc gac cac atg aag        240
Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65              70                  75                  80

cag cac gac ttc ttc aag tcc gcc atg ccc gaa ggc tac gtc cag gag        288
Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

cgc acc atc ttc ttc aag gac gac ggc aac tac aag acc cgc gcc gag        336
Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
                100                 105                 110

gtg aag ttc gag ggc gac acc ctg gtg aac cgc atc gag ctg aag ggc        384
Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
            115                 120                 125

atc gac ttc aag gag gac ggc aac atc ctg ggg cac aac ctg gag tac        432
Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Asn Leu Glu Tyr
        130                 135                 140

aac tac aac agc cac aac gtc tat atc atg gcc gac aag cag aag aac        480
Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145                 150                 155                 160

ggc atc aag gtg aac ttc aag atc cgc cac aac atc gag gac ggc agc        528
Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                 170                 175

gtg cag ctc gcc gac cac tac cag cag aac acc ccc atc ggc gac ggc        576
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190

ccc gtg ctg ctg ccc gac aac cac tac ctg agc acc cag tcc gcc ctg        624
Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
            195                 200                 205

agc aaa gac ccc aac gag aag cgc gat cac atg gtc ctg ctg gag ttc        672
Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
        210                 215                 220

gtg acc gcc gcc ggg atc act ctc ggc atg gac gag ctg tac aag taa       720
Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

```
<210> 12
<211> 239
<212> PRT
<213> Aequorea victoria

<400> 12
Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu
1               5                   10                  15

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
            20                  25                  30
```

```
Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
        35                  40                  45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
    50                  55                  60

Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys
65                  70                  75                  80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
                85                  90                  95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
                100                 105                 110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
        115                 120                 125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Asn Leu Glu Tyr
    130                 135                 140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
145                 150                 155                 160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
                165                 170                 175

Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
            180                 185                 190

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu
        195                 200                 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
    210                 215                 220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys
225                 230                 235
```

<210> 13
<211> 690
<212> DNA
<213> Anemonia majano

<220>
<221> CDS
<222> (1)..(690)
<223> fluorescent protein

<400> 13

```
atg gct ctt tca aac aag ttt atc gga gat gac atg aaa atg acc tac      48
Met Ala Leu Ser Asn Lys Phe Ile Gly Asp Asp Met Lys Met Thr Tyr
1               5                   10                  15

cat atg gat ggc tgt gtc aat ggg cat tac ttt acc gtc aaa ggt gaa      96
His Met Asp Gly Cys Val Asn Gly His Tyr Phe Thr Val Lys Gly Glu
                20                  25                  30
```

```
ggc aac ggg aag cca tac gaa ggg acg cag act tcg act ttt aaa gtc    144
Gly Asn Gly Lys Pro Tyr Glu Gly Thr Gln Thr Ser Thr Phe Lys Val
        35                  40                  45

acc atg gcc aac ggt ggg ccc ctt gca ttc tcc ttt gac ata cta tct    192
Thr Met Ala Asn Gly Gly Pro Leu Ala Phe Ser Phe Asp Ile Leu Ser
    50                  55                  60

aca gtg ttc aaa tat gga aat cga tgc ttt act gcg tat cct acc agt    240
Thr Val Phe Lys Tyr Gly Asn Arg Cys Phe Thr Ala Tyr Pro Thr Ser
65                  70                  75                  80

atg ccc gac tat ttc aaa caa gca ttt cct gac gga atg tca tat gaa    288
Met Pro Asp Tyr Phe Lys Gln Ala Phe Pro Asp Gly Met Ser Tyr Glu
                85                  90                  95

agg act ttt acc tat gaa gat gga gga gtt gct aca gcc agt tgg gaa    336
Arg Thr Phe Thr Tyr Glu Asp Gly Gly Val Ala Thr Ala Ser Trp Glu
            100                 105                 110

ata agc ctt aaa ggc aac tgc ttt gag cac aaa tcc acg ttt cat gga    384
Ile Ser Leu Lys Gly Asn Cys Phe Glu His Lys Ser Thr Phe His Gly
        115                 120                 125

gtg aac ttt cct gct gat gga cct gtg atg gcg aag aag aca act ggt    432
Val Asn Phe Pro Ala Asp Gly Pro Val Met Ala Lys Lys Thr Thr Gly
    130                 135                 140

tgg gac cca tct ttt gag aaa atg act gtc tgc gat gga ata ttg aag    480
Trp Asp Pro Ser Phe Glu Lys Met Thr Val Cys Asp Gly Ile Leu Lys
145                 150                 155                 160

ggt gat gtc acc gcg ttc ctc atg ctg caa gga ggt ggc aat tac aga    528
Gly Asp Val Thr Ala Phe Leu Met Leu Gln Gly Gly Gly Asn Tyr Arg
                165                 170                 175

tgc caa ttc cac act tct tac aag aca aaa aaa ccg gtg acg atg cca    576
Cys Gln Phe His Thr Ser Tyr Lys Thr Lys Lys Pro Val Thr Met Pro
                180                 185                 190

cca aac cat gtg gtg gaa cat cgc att gcg agg acc gac ctt gac aaa    624
Pro Asn His Val Val Glu His Arg Ile Ala Arg Thr Asp Leu Asp Lys
            195                 200                 205

ggt ggc aac agt gtt cag ctg acg gag cac gct gtt gca cat ata acc    672
Gly Gly Asn Ser Val Gln Leu Thr Glu His Ala Val Ala His Ile Thr
        210                 215                 220

tct gtt gtc cct ttc tga                                            690
Ser Val Val Pro Phe
225
```

```
<210> 14
<211> 229
<212> PRT
<213> Anemonia majano
```

<400> 14

```
Met Ala Leu Ser Asn Lys Phe Ile Gly Asp Asp Met Lys Met Thr Tyr
1               5               10              15

His Met Asp Gly Cys Val Asn Gly His Tyr Phe Thr Val Lys Gly Glu
            20              25              30

Gly Asn Gly Lys Pro Tyr Glu Gly Thr Gln Thr Ser Thr Phe Lys Val
        35              40              45

Thr Met Ala Asn Gly Gly Pro Leu Ala Phe Ser Phe Asp Ile Leu Ser
    50              55              60

Thr Val Phe Lys Tyr Gly Asn Arg Cys Phe Thr Ala Tyr Pro Thr Ser
65              70              75              80

Met Pro Asp Tyr Phe Lys Gln Ala Phe Pro Asp Gly Met Ser Tyr Glu
            85              90              95

Arg Thr Phe Thr Tyr Glu Asp Gly Gly Val Ala Thr Ala Ser Trp Glu
        100             105             110

Ile Ser Leu Lys Gly Asn Cys Phe Glu His Lys Ser Thr Phe His Gly
        115             120             125

Val Asn Phe Pro Ala Asp Gly Pro Val Met Ala Lys Lys Thr Thr Gly
    130             135             140

Trp Asp Pro Ser Phe Glu Lys Met Thr Val Cys Asp Gly Ile Leu Lys
145             150             155             160

Gly Asp Val Thr Ala Phe Leu Met Leu Gln Gly Gly Gly Asn Tyr Arg
            165             170             175

Cys Gln Phe His Thr Ser Tyr Lys Thr Lys Lys Pro Val Thr Met Pro
            180             185             190

Pro Asn His Val Val Glu His Arg Ile Ala Arg Thr Asp Leu Asp Lys
        195             200             205

Gly Gly Asn Ser Val Gln Leu Thr Glu His Ala Val Ala His Ile Thr
    210             215             220

Ser Val Val Pro Phe
225
```

<210> 15
<211> 696
<212> DNA
<213> Zoanthus sp.

<220>
<221> CDS
<222> (1)..(696)
<223> fluorescent protein

<400> 15

```
atg gct cag tca aag cac ggt cta aca aaa gaa atg aca atg aaa tac        48
Met Ala Gln Ser Lys His Gly Leu Thr Lys Glu Met Thr Met Lys Tyr
1               5                   10                  15

cgt atg gaa ggg tgc gtc gat gga cat aaa ttt gtg atc acg gga gag        96
Arg Met Glu Gly Cys Val Asp Gly His Lys Phe Val Ile Thr Gly Glu
                20                  25                  30

ggc att gga tat ccg ttc aaa ggg aaa cag gct att aat ctg tgt gtg       144
Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Ala Ile Asn Leu Cys Val
            35                  40                  45

gtc gaa ggt gga cca ttg cca ttt gcc gaa gac ata ttg tca gct gcc       192
Val Glu Gly Gly Pro Leu Pro Phe Ala Glu Asp Ile Leu Ser Ala Ala
        50                  55                  60

ttt aac tac gga aac agg gtt ttc act gaa tat cct caa gac ata gtt       240
Phe Asn Tyr Gly Asn Arg Val Phe Thr Glu Tyr Pro Gln Asp Ile Val
65                  70                  75                  80

gac tat ttc aag aac tcg tgt cct gct gga tat aca tgg gac agg tct       288
Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Asp Arg Ser
                85                  90                  95

ttt ctc ttt gag gat gga gca gtt tgc ata tgt aat gca gat ata aca       336
Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Ala Asp Ile Thr
                100                 105                 110

gtg agt gtt gaa gaa aac tgc atg tat cat gag tcc aaa ttt tat gga       384
Val Ser Val Glu Glu Asn Cys Met Tyr His Glu Ser Lys Phe Tyr Gly
        115                 120                 125

gtg aat ttt cct gct gat gga cct gtg atg aaa aag atg aca gat aac       432
Val Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Asp Asn
        130                 135                 140

tgg gag cca tcc tgc gag aag atc ata cca gta cct aag cag ggg ata       480
Trp Glu Pro Ser Cys Glu Lys Ile Ile Pro Val Pro Lys Gln Gly Ile
145                 150                 155                 160

ttg aaa ggg gat gtc tcc atg tac ctc ctt ctg aag gat ggt ggg cgt       528
Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg
                165                 170                 175

tta cgg tgc caa ttc gac aca gtt tac aaa gca aag tct gtg cca aga       576
Leu Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Arg
                180                 185                 190

aag atg ccg gac tgg cac ttc atc cag cat aag ctc acc cgt gaa gac       624
Lys Met Pro Asp Trp His Phe Ile Gln His Lys Leu Thr Arg Glu Asp
        195                 200                 205

cgc agc gat gct aag aat cag aaa tgg cat ctg aca gaa cat gct att       672
Arg Ser Asp Ala Lys Asn Gln Lys Trp His Leu Thr Glu His Ala Ile
        210                 215                 220
```

```
gca tcc gga tct gca ttg ccc tga                                        696
Ala Ser Gly Ser Ala Leu Pro
225             230
```

<210> 16
<211> 231
<212> PRT
<213> Zoanthus sp.

<400> 16

```
Met Ala Gln Ser Lys His Gly Leu Thr Lys Glu Met Thr Met Lys Tyr
1               5               10              15

Arg Met Glu Gly Cys Val Asp Gly His Lys Phe Val Ile Thr Gly Glu
            20              25              30

Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Ala Ile Asn Leu Cys Val
        35              40              45

Val Glu Gly Gly Pro Leu Pro Phe Ala Glu Asp Ile Leu Ser Ala Ala
    50              55              60

Phe Asn Tyr Gly Asn Arg Val Phe Thr Glu Tyr Pro Gln Asp Ile Val
65              70              75              80

Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Asp Arg Ser
            85              90              95

Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Ala Asp Ile Thr
            100             105             110

Val Ser Val Glu Glu Asn Cys Met Tyr His Glu Ser Lys Phe Tyr Gly
        115             120             125

Val Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Asp Asn
    130             135             140

Trp Glu Pro Ser Cys Glu Lys Ile Ile Pro Val Pro Lys Gln Gly Ile
145             150             155             160

Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg
            165             170             175

Leu Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Arg
            180             185             190

Lys Met Pro Asp Trp His Phe Ile Gln His Lys Leu Thr Arg Glu Asp
        195             200             205

Arg Ser Asp Ala Lys Asn Gln Lys Trp His Leu Thr Glu His Ala Ile
    210             215             220

Ala Ser Gly Ser Ala Leu Pro
225             230
```

<210> 17
<211> 696
<212> DNA
<213> Zoanthus sp.

<220>
<221> CDS
<222> (1)..(696)
<223> fluorescent protein

<400> 17

```
atg gct cat tca aag cac ggt cta aaa gaa gaa atg aca atg aaa tac      48
Met Ala His Ser Lys His Gly Leu Lys Glu Glu Met Thr Met Lys Tyr
1               5                   10                  15

cac atg gaa ggg tgc gtc aac gga cat aaa ttt gtg atc acg ggc gaa      96
His Met Glu Gly Cys Val Asn Gly His Lys Phe Val Ile Thr Gly Glu
                20                  25                  30

ggc att gga tat ccg ttc aaa ggg aaa cag act att aat ctg tgt gtg     144
Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Thr Ile Asn Leu Cys Val
            35                  40                  45

atc gaa ggg gga cca ttg cca ttt tcc gaa gac ata ttg tca gct ggc     192
Ile Glu Gly Gly Pro Leu Pro Phe Ser Glu Asp Ile Leu Ser Ala Gly
        50                  55                  60

ttt aag tac gga gac agg att ttc act gaa tat cct caa gac ata gta     240
Phe Lys Tyr Gly Asp Arg Ile Phe Thr Glu Tyr Pro Gln Asp Ile Val
65                  70                  75                  80

gac tat ttc aag aac tcg tgt cct gct gga tat aca tgg ggc agg tct     288
Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Gly Arg Ser
                85                  90                  95

ttt ctc ttt gag gat gga gca gtc tgc ata tgc aat gta gat ata aca     336
Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Val Asp Ile Thr
                100                 105                 110

gtg agt gtc aaa gaa aac tgc att tat cat aag agc ata ttt aat gga     384
Val Ser Val Lys Glu Asn Cys Ile Tyr His Lys Ser Ile Phe Asn Gly
            115                 120                 125

atg aat ttt cct gct gat gga cct gtg atg aaa aag atg aca act aac     432
Met Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Thr Asn
        130                 135                 140

tgg gaa gca tcc tgc gag aag atc atg cca gta cct aag cag ggg ata     480
Trp Glu Ala Ser Cys Glu Lys Ile Met Pro Val Pro Lys Gln Gly Ile
145                 150                 155                 160

ctg aaa ggg gat gtc tcc atg tac ctc ctt ctg aag gat ggt ggg cgt     528
Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg
                165                 170                 175

tac cgg tgc cag ttc gac aca gtt tac aaa gca aag tct gtg cca agt     576
Tyr Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Ser
                180                 185                 190
```

```
aag atg ccg gag tgg cac ttc atc cag cat aag ctc ctc cgt gaa gac      624
Lys Met Pro Glu Trp His Phe Ile Gln His Lys Leu Leu Arg Glu Asp
        195                 200                 205

cgc agc gat gct aag aat cag aag tgg cag ctg aca gag cat gct att      672
Arg Ser Asp Ala Lys Asn Gln Lys Trp Gln Leu Thr Glu His Ala Ile
        210                 215                 220

gca ttc cct tct gcc ttg gcc tga                                      696
Ala Phe Pro Ser Ala Leu Ala
225                 230
```

<210> 18
<211> 231
<212> PRT
<213> Zoanthus sp.

<400> 18

```
Met Ala His Ser Lys His Gly Leu Lys Glu Glu Met Thr Met Lys Tyr
1               5               10              15

His Met Glu Gly Cys Val Asn Gly His Lys Phe Val Ile Thr Gly Glu
            20              25              30

Gly Ile Gly Tyr Pro Phe Lys Gly Lys Gln Thr Ile Asn Leu Cys Val
            35              40              45

Ile Glu Gly Gly Pro Leu Pro Phe Ser Glu Asp Ile Leu Ser Ala Gly
    50              55              60

Phe Lys Tyr Gly Asp Arg Ile Phe Thr Glu Tyr Pro Gln Asp Ile Val
65              70              75              80

Asp Tyr Phe Lys Asn Ser Cys Pro Ala Gly Tyr Thr Trp Gly Arg Ser
                85              90              95

Phe Leu Phe Glu Asp Gly Ala Val Cys Ile Cys Asn Val Asp Ile Thr
            100             105             110

Val Ser Val Lys Glu Asn Cys Ile Tyr His Lys Ser Ile Phe Asn Gly
            115             120             125

Met Asn Phe Pro Ala Asp Gly Pro Val Met Lys Lys Met Thr Thr Asn
    130             135             140

Trp Glu Ala Ser Cys Glu Lys Ile Met Pro Val Pro Lys Gln Gly Ile
145             150             155             160

Leu Lys Gly Asp Val Ser Met Tyr Leu Leu Leu Lys Asp Gly Gly Arg
            165             170             175

Tyr Arg Cys Gln Phe Asp Thr Val Tyr Lys Ala Lys Ser Val Pro Ser
            180             185             190

Lys Met Pro Glu Trp His Phe Ile Gln His Lys Leu Leu Arg Glu Asp
        195             200             205
```

Arg Ser Asp Ala Lys Asn Gln Lys Trp Gln Leu Thr Glu His Ala Ile
    210   215   220

Ala Phe Pro Ser Ala Leu Ala
225   230

<210> 19
<211> 699
<212> DNA
<213> Discosoma striata

<220>
<221> CDS
<222> (1)..(699)
<223> fluorescent protein

<400> 19
atg agt tgt tcc aag agt gtg atc aag gaa gaa atg ttg atc gat ctt   48
Met Ser Cys Ser Lys Ser Val Ile Lys Glu Glu Met Leu Ile Asp Leu
1   5   10   15

cat ctg gaa gga acg ttc aat ggg cac tac ttt gaa ata aaa ggc aaa   96
His Leu Glu Gly Thr Phe Asn Gly His Tyr Phe Glu Ile Lys Gly Lys
                20   25   30

gga aaa gga cag cct aat gaa ggc acc aat acc gtc acg ctc gag gtt   144
Gly Lys Gly Gln Pro Asn Glu Gly Thr Asn Thr Val Thr Leu Glu Val   .
            35   40   45

acc aag ggt gga cct ctg cca ttt ggt tgg cat att ttg tgc cca caa   192
Thr Lys Gly Gly Pro Leu Pro Phe Gly Trp His Ile Leu Cys Pro Gln
        50   55   60

ttt cag tat gga aac aag gca ttt gtc cac cac cct gac aac ata cat   240
Phe Gln Tyr Gly Asn Lys Ala Phe Val His His Pro Asp Asn Ile His
65   70   75   80

gat tat cta aag ctg tca ttt ccg gag gga tat aca tgg gaa cgg tcc   288
Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ser
                85   90   95

atg cac ttt gaa gac ggt ggc ttg tgt tgt atc acc aat gat atc agt   336
Met His Phe Glu Asp Gly Gly Leu Cys Cys Ile Thr Asn Asp Ile Ser
                100   105   110

ttg aca ggc aac tgt ttc tac tac gac atc aag ttc act ggc ttg aac   384
Leu Thr Gly Asn Cys Phe Tyr Tyr Asp Ile Lys Phe Thr Gly Leu Asn
            115   120   125

ttt cct cca aat gga ccc gtt gtg cag aag aag aca act ggc tgg gaa   432
Phe Pro Pro Asn Gly Pro Val Val Gln Lys Lys Thr Thr Gly Trp Glu
        130   135   140

ccg agc act gag cgt ttg tat cct cgt gat ggt gtg ttg ata gga gac   480
Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Ile Gly Asp
145   150   155   160

```
atc cat cat gct ctg aca gtt gaa gga ggt ggt cat tac gca tgt gac       528
Ile His His Ala Leu Thr Val Glu Gly Gly Gly His Tyr Ala Cys Asp
                165                 170                 175

att aaa act gtt tac agg gcc aag aag gcc gcc ttg aag atg cca ggg       576
Ile Lys Thr Val Tyr Arg Ala Lys Lys Ala Ala Leu Lys Met Pro Gly
                180                 185                 190

tat cac tat gtt gac acc aaa ctg gtt ata tgg aac aac gac aaa gaa       624
Tyr His Tyr Val Asp Thr Lys Leu Val Ile Trp Asn Asn Asp Lys Glu
                195                 200                 205

ttc atg aaa gtt gag gag cat gaa atc gcc gtt gca cgc cac cat ccg       672
Phe Met Lys Val Glu Glu His Glu Ile Ala Val Ala Arg His His Pro
    210                 215                 220

ttc tat gag cca aag aag gat aag taa                                   699
Phe Tyr Glu Pro Lys Lys Asp Lys
225                 230
```

<210> 20
<211> 232
<212> PRT
<213> Discosoma striata

<400> 20

```
Met Ser Cys Ser Lys Ser Val Ile Lys Glu Glu Met Leu Ile Asp Leu
1               5                   10                  15

His Leu Glu Gly Thr Phe Asn Gly His Tyr Phe Glu Ile Lys Gly Lys
            20                  25                  30

Gly Lys Gly Gln Pro Asn Glu Gly Thr Asn Thr Val Thr Leu Glu Val
        35                  40                  45

Thr Lys Gly Gly Pro Leu Pro Phe Gly Trp His Ile Leu Cys Pro Gln
    50                  55                  60

Phe Gln Tyr Gly Asn Lys Ala Phe Val His His Pro Asp Asn Ile His
65                  70                  75                  80

Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Tyr Thr Trp Glu Arg Ser
            85                  90                  95

Met His Phe Glu Asp Gly Gly Leu Cys Cys Ile Thr Asn Asp Ile Ser
            100                 105                 110

Leu Thr Gly Asn Cys Phe Tyr Tyr Asp Ile Lys Phe Thr Gly Leu Asn
            115                 120                 125

Phe Pro Pro Asn Gly Pro Val Val Gln Lys Lys Thr Thr Gly Trp Glu
    130                 135                 140

Pro Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Ile Gly Asp
145                 150                 155                 160
```

```
Ile His His Ala Leu Thr Val Glu Gly Gly Gly His Tyr Ala Cys Asp
                165             170             175

Ile Lys Thr Val Tyr Arg Ala Lys Lys Ala Ala Leu Lys Met Pro Gly
            180             185             190

Tyr His Tyr Val Asp Thr Lys Leu Val Ile Trp Asn Asn Asp Lys Glu
            195             200             205

Phe Met Lys Val Glu Glu His Glu Ile Ala Val Ala Arg His His Pro
    210             215             220

Phe Tyr Glu Pro Lys Lys Asp Lys
225             230
```

<210> 21
<211> 678
<212> DNA
<213> Discosoma sp.

<220>
<221> CDS
<222> (1)..(678)
<223> fluorescent protein

<400> 21
```
atg agg tct tcc aag aat gtt atc aag gag ttc atg agg ttt aag gtt          48
Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val
1                   5                   10                  15

cgc atg gaa gga acg gtc aat ggg cac gag ttt gaa ata gaa ggc gaa          96
Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu
                20                  25                  30

gga gag ggg agg cca tac gaa ggc cac aat acc gta aag ctt aag gta         144
Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val
            35                  40                  45

acc aag ggg gga cct ttg cca ttt gct tgg gat att ttg tca cca caa         192
Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln
    50                  55                  60

ttt cag tat gga agc aag gta tat gtc aag cac cct gcc gac ata cca         240
Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro
65                  70                  75                  80

gac tat aaa aag ctg tca ttt cct gaa gga ttt aaa tgg gaa agg gtc         288
Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val
                85                  90                  95

atg aac ttt gaa gac ggt ggc gtc gtt act gta acc cag gat tcc agt         336
Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser
                100                 105                 110

ttg cag gat ggc tgt ttc atc tac aag gtc aag ttc att ggc gtg aac         384
Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn
            115                 120                 125
```

```
ttt cct tcc gat gga cct gtt atg caa aag aag aca atg ggc tgg gaa    432
Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu
    130             135             140

gcc agc act gag cgt ttg tat cct cgt gat ggc gtg ttg aaa gga gag    480
Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu
145             150             155             160

att cat aag gct ctg aag ctg aaa gac ggt ggt cat tac cta gtt gaa    528
Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu
                165             170             175

ttc aaa agt att tac atg gca aag aag cct gtg cag cta cca ggg tac    576
Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr
                180             185             190

tac tat gtt gac tcc aaa ctg gat ata aca agc cac aac gaa gac tat    624
Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr
            195             200             205

aca atc gtt gag cag tat gaa aga acc gag gga cgc cac cat ctg ttc    672
Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe
    210             215             220

ctt taa                                                            678
Leu
225
```

<210> 22
<211> 225
<212> PRT
<213> Discosoma sp.

<400> 22

```
Met Arg Ser Ser Lys Asn Val Ile Lys Glu Phe Met Arg Phe Lys Val
1               5               10              15

Arg Met Glu Gly Thr Val Asn Gly His Glu Phe Glu Ile Glu Gly Glu
            20              25              30

Gly Glu Gly Arg Pro Tyr Glu Gly His Asn Thr Val Lys Leu Lys Val
        35              40              45

Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln
    50              55              60

Phe Gln Tyr Gly Ser Lys Val Tyr Val Lys His Pro Ala Asp Ile Pro
65              70              75              80

Asp Tyr Lys Lys Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val
            85              90              95

Met Asn Phe Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser
            100             105             110
```

```
Leu Gln Asp Gly Cys Phe Ile Tyr Lys Val Lys Phe Ile Gly Val Asn
        115                 120                 125

Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu
    130                 135                 140

Ala Ser Thr Glu Arg Leu Tyr Pro Arg Asp Gly Val Leu Lys Gly Glu
145                 150                 155                 160

Ile His Lys Ala Leu Lys Leu Lys Asp Gly Gly His Tyr Leu Val Glu
                165                 170                 175

Phe Lys Ser Ile Tyr Met Ala Lys Lys Pro Val Gln Leu Pro Gly Tyr
                180                 185                 190

Tyr Tyr Val Asp Ser Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr
        195                 200                 205

Thr Ile Val Glu Gln Tyr Glu Arg Thr Glu Gly Arg His His Leu Phe
    210                 215                 220

Leu
225


<210> 23
<211> 801
<212> DNA
<213> Clavularia sp.

<220>
<221> CDS
<222> (1)..(801)
<223> fluorescent protein

<400> 23
atg aag tgt aaa ttt gtg ttc tgc ctg tcc ttc ttg gtc ctc gcc atc        48
Met Lys Cys Lys Phe Val Phe Cys Leu Ser Phe Leu Val Leu Ala Ile
1               5                   10                  15

aca aac gcg aac att ttt ttg aga aac gag gct gac tta gaa gag aag        96
Thr Asn Ala Asn Ile Phe Leu Arg Asn Glu Ala Asp Leu Glu Glu Lys
                20                  25                  30

aca ttg aga ata cca aaa gct cta acc acc atg ggt gtg att aaa cca        144
Thr Leu Arg Ile Pro Lys Ala Leu Thr Thr Met Gly Val Ile Lys Pro
            35                  40                  45

gac atg aag att aag ctg aag atg gaa gga aat gta aac ggg cat gct        192
Asp Met Lys Ile Lys Leu Lys Met Glu Gly Asn Val Asn Gly His Ala
    50                  55                  60

ttt gtg atc gaa gga gaa gga gaa gga aag cct tac gat ggg aca cac        240
Phe Val Ile Glu Gly Glu Gly Glu Gly Lys Pro Tyr Asp Gly Thr His
65                  70                  75                  80
```

```
act tta aac ctg gaa gtg aag gaa ggt gcg cct ctg cct ttt tct tac          288
Thr Leu Asn Leu Glu Val Lys Glu Gly Ala Pro Leu Pro Phe Ser Tyr
                85                  90                  95

gat atc ttg tca aac gcg ttc cag tac gga aac aga gca ttg aca aaa          336
Asp Ile Leu Ser Asn Ala Phe Gln Tyr Gly Asn Arg Ala Leu Thr Lys
                100                 105                 110

tac cca gac gat ata gca gac tat ttc aag cag tcg ttt ccc gag gga          384
Tyr Pro Asp Asp Ile Ala Asp Tyr Phe Lys Gln Ser Phe Pro Glu Gly
            115                 120                 125

tat tcc tgg gaa aga acc atg act ttt gaa gac aaa ggc att gtc aaa          432
Tyr Ser Trp Glu Arg Thr Met Thr Phe Glu Asp Lys Gly Ile Val Lys
    130                 135                 140

gtg aaa agt gac ata agc atg gag gaa gac tcc ttt atc tat gaa att          480
Val Lys Ser Asp Ile Ser Met Glu Glu Asp Ser Phe Ile Tyr Glu Ile
145                 150                 155                 160

cgt ttt gat ggg atg aac ttt cct ccc aat ggt ccg gtt atg cag aaa          528
Arg Phe Asp Gly Met Asn Phe Pro Pro Asn Gly Pro Val Met Gln Lys
                165                 170                 175

aaa act ttg aag tgg gaa cca tcc act gag att atg tac gtg cgt gat          576
Lys Thr Leu Lys Trp Glu Pro Ser Thr Glu Ile Met Tyr Val Arg Asp
            180                 185                 190

gga gtg ctg gtc gga gat att agc cat tct ctg ttg ctg gag gga ggt          624
Gly Val Leu Val Gly Asp Ile Ser His Ser Leu Leu Leu Glu Gly Gly
            195                 200                 205

ggc cat tac cga tgt gac ttc aaa agt att tac aaa gca aaa aaa gtt          672
Gly His Tyr Arg Cys Asp Phe Lys Ser Ile Tyr Lys Ala Lys Lys Val
    210                 215                 220

gtc aaa ttg cca gac tat cac ttt gtg gac cat cgc att gag atc ttg          720
Val Lys Leu Pro Asp Tyr His Phe Val Asp His Arg Ile Glu Ile Leu
225                 230                 235                 240

aac cat gac aag gat tac aac aaa gta acg ctg tat gag aat gca gtt          768
Asn His Asp Lys Asp Tyr Asn Lys Val Thr Leu Tyr Glu Asn Ala Val
                245                 250                 255

gct cgc tat tct ttg ctg cca agt cag gcc tag                              801
Ala Arg Tyr Ser Leu Leu Pro Ser Gln Ala
            260                 265


<210> 24
<211> 266
<212> PRT
<213> Clavularia sp.

<400> 24
Met Lys Cys Lys Phe Val Phe Cys Leu Ser Phe Leu Val Leu Ala Ile
1               5                   10                  15
```

Thr Asn Ala Asn Ile Phe Leu Arg Asn Glu Ala Asp Leu Glu Glu Lys
            20   25   30

Thr Leu Arg Ile Pro Lys Ala Leu Thr Thr Met Gly Val Ile Lys Pro
            35   40   45

Asp Met Lys Ile Lys Leu Lys Met Glu Gly Asn Val Asn Gly His Ala
        50   55   60

Phe Val Ile Glu Gly Glu Gly Glu Gly Lys Pro Tyr Asp Gly Thr His
65   70   75   80

Thr Leu Asn Leu Glu Val Lys Glu Gly Ala Pro Leu Pro Phe Ser Tyr
                85   90   95

Asp Ile Leu Ser Asn Ala Phe Gln Tyr Gly Asn Arg Ala Leu Thr Lys
            100   105   110

Tyr Pro Asp Asp Ile Ala Asp Tyr Phe Lys Gln Ser Phe Pro Glu Gly
        115   120   125

Tyr Ser Trp Glu Arg Thr Met Thr Phe Glu Asp Lys Gly Ile Val Lys
    130   135   140

Val Lys Ser Asp Ile Ser Met Glu Glu Asp Ser Phe Ile Tyr Glu Ile
145   150   155   160

Arg Phe Asp Gly Met Asn Phe Pro Pro Asn Gly Pro Val Met Gln Lys
                165   170   175

Lys Thr Leu Lys Trp Glu Pro Ser Thr Glu Ile Met Tyr Val Arg Asp
            180   185   190

Gly Val Leu Val Gly Asp Ile Ser His Ser Leu Leu Leu Glu Gly Gly
            195   200   205

Gly His Tyr Arg Cys Asp Phe Lys Ser Ile Tyr Lys Ala Lys Lys Val
    210   215   220

Val Lys Leu Pro Asp Tyr His Phe Val Asp His Arg Ile Glu Ile Leu
225   230   235   240

Asn His Asp Lys Asp Tyr Asn Lys Val Thr Leu Tyr Glu Asn Ala Val
                245   250   255

Ala Arg Tyr Ser Leu Leu Pro Ser Gln Ala
            260   265

<210> 25
<211> 881
<212> PRT
<213> Saccharomyces cerevisiae

<400> 25
Met Lys Leu Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu
1   5   10   15

Lys Lys Leu Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu
        20   25   30

Lys Asn Asn Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro
        35   40   45

Leu Thr Arg Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu
    50   55   60

Glu Gln Leu Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile
65   70   75   80

Leu Lys Met Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu
            85   90   95

Phe Val Gln Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala
            100   105   110

Ser Val Glu Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser
        115   120   125

Ala Thr Ser Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu
    130   135   140

Thr Val Ser Ile Asp Ser Ala Ala His His Asp Asn Ser Thr Ile Pro
145   150   155   160

Leu Asp Phe Met Pro Arg Asp Ala Leu His Gly Phe Asp Trp Ser Glu
            165   170   175

Glu Asp Asp Met Ser Asp Gly Leu Pro Phe Leu Lys Thr Asp Pro Asn
            180   185   190

Asn Asn Gly Phe Phe Gly Asp Gly Ser Leu Leu Cys Ile Leu Arg Ser
        195   200   205

Ile Gly Phe Lys Pro Glu Asn Tyr Thr Asn Ser Asn Val Asn Arg Leu
    210   215   220

Pro Thr Met Ile Thr Asp Arg Tyr Thr Leu Ala Ser Arg Ser Thr Thr
225   230   235   240

Ser Arg Leu Leu Gln Ser Tyr Leu Asn Asn Phe His Pro Tyr Cys Pro
            245   250   255

Ile Val His Ser Pro Thr Leu Met Met Leu Tyr Asn Asn Gln Ile Glu
        260   265   270

Ile Ala Ser Lys Asp Gln Trp Gln Ile Leu Phe Asn Cys Ile Leu Ala
        275   280   285

Ile Gly Ala Trp Cys Ile Glu Gly Glu Ser Thr Asp Ile Asp Val Phe
    290   295   300

Tyr Tyr Gln Asn Ala Lys Ser His Leu Thr Ser Lys Val Phe Glu Ser
305   310   315   320

Gly Ser Ile Ile Leu Val Thr Ala Leu His Leu Leu Ser Arg Tyr Thr
                325   330   335

Gln Trp Arg Gln Lys Thr Asn Thr Ser Tyr Asn Phe His Ser Phe Ser
            340   345   350

Ile Arg Met Ala Ile Ser Leu Gly Leu Asn Arg Asp Leu Pro Ser Ser
        355   360   365

Phe Ser Asp Ser Ser Ile Leu Glu Gln Arg Arg Arg Ile Trp Trp Ser
    370   375   380

Val Tyr Ser Trp Glu Ile Gln Leu Ser Leu Leu Tyr Gly Arg Ser Ile
385   390   395   400

Gln Leu Ser Gln Asn Thr Ile Ser Phe Pro Ser Ser Val Asp Asp Val
            405   410   415

Gln Arg Thr Thr Thr Gly Pro Thr Ile Tyr His Gly Ile Ile Glu Thr
        420   425   430

Ala Arg Leu Leu Gln Val Phe Thr Lys Ile Tyr Glu Leu Asp Lys Thr
        435   440   445

Val Thr Ala Glu Lys Ser Pro Ile Cys Ala Lys Lys Cys Leu Met Ile
    450   455   460

Cys Asn Glu Ile Glu Glu Val Ser Arg Gln Ala Pro Lys Phe Leu Gln
465   470   475   480

Met Asp Ile Ser Thr Thr Ala Leu Thr Asn Leu Leu Lys Glu His Pro
            485   490   495

Trp Leu Ser Phe Thr Arg Phe Glu Leu Lys Trp Lys Gln Leu Ser Leu
            500   505   510

Ile Ile Tyr Val Leu Arg Asp Phe Phe Thr Asn Phe Thr Gln Lys Lys
        515   520   525

Ser Gln Leu Glu Gln Asp Gln Asn Asp His Gln Ser Tyr Glu Val Lys
    530   535   540

Arg Cys Ser Ile Met Leu Ser Asp Ala Ala Gln Arg Thr Val Met Ser
545   550   555   560

Val Ser Ser Tyr Met Asp Asn His Asn Val Thr Pro Tyr Phe Ala Trp
            565   570   575

Asn Cys Ser Tyr Tyr Leu Phe Asn Ala Val Leu Val Pro Ile Lys Thr
        580   585   590

Leu Leu Ser Asn Ser Lys Ser Asn Ala Glu Asn Asn Glu Thr Ala Gln
        595   600   605

Leu Leu Gln Gln Ile Asn Thr Val Leu Met Leu Leu Lys Lys Leu Ala
    610   615   620

Thr Phe Lys Ile Gln Thr Cys Glu Lys Tyr Ile Gln Val Leu Glu Glu
625   630   635   640

Val Cys Ala Pro Phe Leu Leu Ser Gln Cys Ala Ile Pro Leu Pro His
                645   650   655

Ile Ser Tyr Asn Asn Ser Asn Gly Ser Ala Ile Lys Asn Ile Val Gly
            660   665   670

Ser Ala Thr Ile Ala Gln Tyr Pro Thr Leu Pro Glu Glu Asn Val Asn
        675   680   685

Asn Ile Ser Val Lys Tyr Val Ser Pro Gly Ser Val Gly Pro Ser Pro
    690   695   700

Val Pro Leu Lys Ser Gly Ala Ser Phe Ser Asp Leu Val Lys Leu Leu
705   710   715   720

Ser Asn Arg Pro Pro Ser Arg Asn Ser Pro Val Thr Ile Pro Arg Ser
            725   730   735

Thr Pro Ser His Arg Ser Val Thr Pro Phe Leu Gly Gln Gln Gln Gln
            740   745   750

Leu Gln Ser Leu Val Pro Leu Thr Pro Ser Ala Leu Phe Gly Gly Ala
        755   760   765

Asn Phe Asn Gln Ser Gly Asn Ile Ala Asp Ser Ser Leu Ser Phe Thr
    770   775   780

Phe Thr Asn Ser Ser Asn Gly Pro Asn Leu Ile Thr Thr Gln Thr Asn
785   790   795   800

Ser Gln Ala Leu Ser Gln Pro Ile Ala Ser Ser Asn Val His Asp Asn
            805   810   815

Phe Met Asn Asn Glu Ile Thr Ala Ser Lys Ile Asp Asp Gly Asn Asn
        820   825   830

Ser Lys Pro Leu Ser Pro Gly Trp Thr Asp Gln Thr Ala Tyr Asn Ala
        835   840   845

Phe Gly Ile Thr Thr Gly Met Phe Asn Thr Thr Thr Met Asp Asp Val
        850   855   860

Tyr Asn Tyr Leu Phe Asp Asp Glu Asp Thr Pro Pro Asn Pro Lys Lys
865   870   875   880

Glu

<210> 26
<211> 202
<212> PRT
<213> Saccharomyces cerevisiae

<400> 26
Met Lys Ala Leu Thr Ala Arg Gln Gln Glu Val Phe Asp Leu Ile Arg
1   5   10   15

Asp His Ile Ser Gln Thr Gly Met Pro Pro Thr Arg Ala Glu Ile Ala
            20   25   30

Gln Arg Leu Gly Phe Arg Ser Pro Asn Ala Ala Glu Glu His Leu Lys
        35   40   45

Ala Leu Ala Arg Lys Gly Val Ile Glu Ile Val Ser Gly Ala Ser Arg
    50   55   60

Gly Ile Arg Leu Leu Gln Glu Glu Glu Gly Leu Pro Leu Val Gly
65   70   75   80

Arg Val Ala Ala Gly Glu Pro Leu Leu Ala Gln Gln His Ile Glu Gly
            85   90   95

His Tyr Gln Val Asp Pro Ser Leu Phe Lys Pro Asn Ala Asp Phe Leu
            100   105   110

Leu Arg Val Ser Gly Met Ser Met Lys Asp Ile Gly Ile Met Asp Gly
        115   120   125

Asp Leu Leu Ala Val His Lys Thr Gln Asp Val Arg Asn Gly Gln Val
    130   135   140

Val Val Ala Arg Ile Asp Asp Glu Val Thr Val Lys Arg Leu Lys Lys
145   150   155   160

Gln Gly Asn Lys Val Glu Leu Leu Pro Glu Asn Ser Glu Phe Lys Pro
            165   170   175

Ile Val Val Asp Leu Arg Gln Gln Ser Phe Thr Ile Glu Gly Leu Ala
            180   185   190

Val Gly Val Ile Arg Asn Gly Asp Trp Leu
        195   200

<210> 27
<211> 17
<212> DNA
<213> Unknown Organism

<220>
<223> Description of Unknown Organism: Nucleotide sequence that
      is bound by GAL4

<220>
<221> modified_base
<222> (4)..(4)
<223> any purine

<220>
<221> modified_base
<222> (5)..(5)
<223> any nucleotide

<220>
<221> modified_base
<222> (6)..(6)
<223> any nucleotide

<220>
<221> modified_base
<222> (7)..(7)
<223> any purine

<220>
<221> modified_base
<222> (9)..(9)
<223> any pyrimidine

<220>
<221> modified_base
<222> (10)..(10)
<223> any nucleotide

<220>
<221> modified_base
<222> (11)..(11)
<223> any pyrimidine

<220>
<221> modified_base
<222> (12)..(12)
<223> any nucleotide

<220>
<221> modified_base
<222> (14)..(14)
<223> any nucleotide

<400> 27
cggnnnncnn nncnccg   17

<210> 28
<211> 14
<212> DNA
<213> Unknown Organism

<220>
<223> Description of Unknown Organism: Nucleotide sequence
      that binds LexA

<220>
<221> modified_base
<222> (6)..(9)
<223> any nucleotide

<400> 28
cgaacnnnng ttcg   14

<210> 29
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 29
ctagacgttg acattgatta ttgac   25

<210> 30
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 30
tctagagcca gtaagcagtg ggttc   25

<210> 31
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 31
gcggccgccc attgcattgg gattggtc   28

<210> 32
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 32
gcggccgccc tccaagtctt tatctg   26

<210> 33
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 33
ggtaccggct cgtagagaaa tatgaacc   28

<210> 34
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 34
ctcgagagag actgaatttc cctttt   26

<210> 35
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 35
gcggccgcga aggaagatag ccaaggctta   30

<210> 36
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 36
gcggccgctt aaagatgcca cactgctggt ct   32

<210> 37
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 37
gtcgacacct gtcggctact gggata   26

<210> 38
<211> 26
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 38
gatatcgtgg gactctgttt ccagca   26

<210> 39
<211> 6795
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
      DNA vector sequence

<220>
<221> modified_base
<222> (892)..(892)
<223> unknown nucleotide

<400> 39
aagtccggca cctcgtgcac gcggatttcg gctccaacaa tgtcctgacg gacaatggcc   60
gcataacagc ggtcattgac tggagcgagg cgatgttcgg ggattcccaa tacgaggtcg   120
ccaacatctt cttctggagg ccgtggttgg cttgtatgga gcagcagacg cgctacttcg   180
agcggaggca tccggagctt gcaggatcgc cgcggctccg ggcgtatatg ctccgcattg   240
gtcttgacca actctatcag agcttggttg acggcaattt cgatgatgca gcttggcgc   300
agggtcgatg cgacgcaatc gtccgatccg gagccgggac tgtcgggcgt acacaaatcg   360
cccgcagaag cgcggccgtc tggaccgatg gctgtgtaga agtactcgcc gatagtggaa   420
accgacgccc cagcactcgt ccgagggcaa aggaatgcct gagaaaggaa gtgagctgta   480
aaggctgagc tctctctctg acgtatgtag cctctggtta gcttcgtcac tcactgttct   540
tgactcagca tggcaatctg atgaaatccc agctgtaagt ctgcataaat tgatgatcta   600
ttaaacaata aagatgtcca ctaaaatgga agttttttac tgtcatactt tgttaagaag   660
ggtgagaaca gagtacctac attttgaatg gaaggattgg agctacgggg gtgggggtgg   720
gggtgggatt agataaatgc ctgctcttta ctgaaggctc tttactattg ctttatgata   780
atgtttcata gttggatatc ataatttaaa caagcaaaac caaattaagg gccagctcat   840
tcctccactc atgatctata gatctataga tctctcgtgg gatcattgtt tntctgatcc   900
actggaagct tatcgatacc gtcgacctcg aggggggggcc cggtacccag cttttgttcc   960
ctttagtgag ggttaattgc gcgcttggcg taatcatggt catagctgtt tcctgtgtga   1020
aattgttatc cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc   1080
tggggtgcct aatgagtgag ctaactcaca ttaattgcgt tgcgctcact gcccgctttc   1140
cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc   1200
ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt   1260
cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca   1320
ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa   1380
aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca tcacaaaaat   1440
cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc   1500
cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc   1560
gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt   1620
tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aaccccccgt tcagcccgac   1680
cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg   1740
ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca   1800
gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt tggtatctgc   1860
gctctgctga gccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa   1920
accaccgctg gtagcggtgg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa   1980
ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac   2040

```
tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatccttta    2100
aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt    2160
taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata    2220
gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc    2280
agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc agcaataaac    2340
cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag    2400
tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac    2460
gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc    2520
agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgtg caaaaaagcg    2580
gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc    2640
atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct    2700
gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc    2760
tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgctc    2820
atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc    2880
agttcgatgt aacccactcg tgcacccaac tgatcttcag catcttttac tttcaccagc    2940
gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca    3000
cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggt    3060
tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca ataggggtt     3120
ccgcgcacat ttccccgaaa agtgccacct aaattgtaag cgttaatatt ttgttaaaat    3180
tcgcgttaaa tttttgttaa atcagctcat tttttaacca ataggccgaa atcggcaaaa    3240
tcccttataa atcaaaagaa tagaccgaga tagggttgag tgttgttcca gtttggaaca    3300
agagtccact attaaagaac gtggactcca acgtcaaagg gcgaaaaacc gtctatcagg    3360
gcgatggccc actacgtgaa ccatcaccct aatcaagttt tttggggtcg aggtgccgta    3420
aagcactaaa tcggaaccct aaagggagcc cccgatttag agcttgacgg ggaaagccgg    3480
cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc gggcgctagg cgctggcaa     3540
gtgtagcggt cacgctgcgc gtaaccacca cacccgccgc gcttaatgcg ccgctacagg    3600
gcgcgtccca ttcgccattc aggctgcgca actgttggga agggcgatcg gtgcgggcct    3660
cttcgctatt acgccagctg gcgaaagggg gatgtgctgc aaggcgatta agttgggtaa    3720
cgccagggtt ttcccagtca cgacgttgta aaacgacggc cagtgagcgc gcgtaatacg    3780
actcactata gggcgaattg gagctccacc gcggtggcgg ccgctccatg cggccacgac    3840
atgataagat acattgatga gtttggacaa accacaacta gaatgcagtg aaaaaaatgc    3900
tttatttgtg aaatttgtga tgctattgct ttatttgtaa ccattataag ctgcaataaa    3960
caagttcttg tacagctcgt ccatgccgag agtgatcccg gcggcggtca cgaactccag    4020
caggaccatg tgatcgcgct tctcgttggg gtctttgctc agggcggact ggtagctcag    4080
gtagtggttg tcgggcagca gcacggggcc gtcgccgatg ggggtgttct gctggtagtg    4140
gtcggcgagc tgcacgctgc cgtcctcgat gttgtggcgg atcttgaagt tcaccttgat    4200
gccgttcttc tgcttgtcgg ccatgatata gacgttgtgg ctgttgtagt tgtactccag    4260
cttgtgcccc aggatgttgc cgtcctcctt gaagtcgatg cccttcagct cgatgcggtt    4320
caccagggtg tcgccctcga acttcacctc ggcgcgggtc ttgtagttgc cgtcgtcctt    4380
gaagaagatg gtgcgctcct ggacgtagcc ttcgggcatg gcggacttga agaagtcgtg    4440
ctgcttcatg tggtcggggt agcgggcgaa gcactgcagg ccgtagccga aggtggtcac    4500
gagggtgggc cagggcacgg gcagcttgcc ggtggtgcag atgaacttca gggtcagctt    4560
gccgtaggtg gcatcgccct cgccctcgcc ggacacgctg aacttgtggc cgtttacgtc    4620
gccgtccagc tcgaccagga tgggcaccac cccggtgaac agctcctcgc ccttgctcac    4680
catggtggcg ttagcttgat ttgacagtgg ctggggggttg cgccgccggg ttttatagga    4740
agccacagcg gccactcgag ccataaaagg caactttagg aacggcgggg ggtgattgga    4800
ttcgagtcgt ttattcaccg gccttccgc acagtgcagc attttttac ccctctccc       4860
ctcctttgc gggggaaaaa aaaaaaaaa aaaaaaaag gagaagagaa aaaagcgag         4920
cgagagagaa agcgagattg aggaagagga tgaagagttt ggcgatgggt gctggttccg     4980
taggcccaga tggacaagaa tagcccccgc ccttgcggac agtatcccat tcagtgactc     5040
agatcagatc aagcggccgc tctagaacta gtggatccgg ataagccagt aagcagtggg     5100
ttctctagtt agccagagag ctctgcttat atagacctcc caccgtacac gcctaccgcc     5160
catttgcgtc aatgggcgg agttgttacg acattttgga aagtcccgtt gattttggtg     5220
ccaaaacaaa ctcccattga cgtcaatggg gtggagactt ggaaatcccc gtgagtcaaa     5280
ccgctatcca cgcccattga tgtactgcca aaaccgcatc accatggtaa tagcgatgac     5340
taatacgtag atgtactgcc aagtaggaaa gtcccataag gtcatgtact gggcataatg     5400
ccaggcgggc catttaccgt cattgacgtc aatagggggc gtacttggca tatgatacac     5460
ttgatgtact gccaagtggg cagtttaccg taaatactcc acccattgac gtcaatggaa     5520
```

```
agtccctatt ggcgttacta tgggaacata cgtcattatt gacgtcaatg ggcgggggtc   5580
gttgggcggt cagccaggcg ggccatttac cgtaagttat gtaacgcgga actccatata   5640
tgggctatga actaatgacc ccgtaattga ttactattaa taactagtca ataatcaatg   5700
tcaaccccgg gctgcaggaa ttctaccggg taggggaggc gcttttccca aggcagtctg   5760
gagcatgcgc tttagcagcc ccgctggcac ttggcgcatc acaagtggcc tctggcctcg   5820
cacacattcc acatccaccg gtagcgccaa ccggctccct tctttggtgg ccccttcgcg   5880
ccaccttcta ctcctcccct agtcaggaag ttccccccccg ccccgcagct cgcgtcgtgc   5940
aggacgtgac aaatggaagt agcacgtctc actagtctcg tgcagatgga caagcaccgc   6000
tgagcaatgg aagcgggtag gcctttgggg cagcggccaa tagcagcttg gctccttcgc   6060
tttctgggct cagaggctgg gaaggggtgg gtccggggggc gggctcaggg gcgggctcag   6120
gggcgggggcg ggcgcgaagg tcctccggac ccggcattct gcacgcttca aaagcgcacg   6180
tctgccgcgc tgttctcctc ttcctcatct ccgggccttc gacctgcatg aaaaagcctg   6240
aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa gttcgacagc gtctccgacc   6300
tgatgcagct ctcggagggc gaagaatctc gtgctttcag cttcgatgta ggagggcgtg   6360
gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt tatgtttatc   6420
ggcactttgc atcggccgcg ctcccgattc cggaagtgct tgacattggg gaattcagcg   6480
agagcctgac ctattgcatc tcccgccgtg cacagggtgt cacgttgcaa gacctgcctg   6540
aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg atcgctgcgg   6600
ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc ggtcaataca   6660
ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac tggcaaactg   6720
tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg atgctttggg   6780
ccgaggactg ccccg                                                     6795
```

<210> 40
<211> 5821
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
      DNA vector sequence

<400> 40
```
ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac     60
tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatccttta    120
aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt    180
taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata    240
gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc    300
agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc agcaataaac    360
cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag    420
tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac    480
gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc    540
agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgtg caaaaaagcg    600
gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc    660
atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct    720
gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc    780
tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgctc    840
atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc    900
agttcgatgt aacccactcg tgcacccaac tgatcttcag catcttttac tttcaccagc    960
gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca   1020
cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggt   1080
tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt   1140
ccgcgcacat ttccccgaaa agtgccacct gacgtcgacg gatcgggaga tctcccgatc   1200
ccctatggtg cactctcagt acaatctgct ctgatccgc atagttaagc cagtatctgc     1260
tccctgcttg tgtgttggag gtcgctgagt agtgcgcgag caaaatttaa gctacaacaa   1320
ggcaaggctt gaccgacaat tgcatgaaga atctgcttag ggttaggcgt tttgcgctgc   1380
ttcgcgatgt acgggccaga tatacgcgtt gacattgatt attgactagt tattaatagt   1440
```

```
aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta   1500
cggtaaatgg cccgcctggc tgaccgccca acgacccccg cccattgacg tcaataatga   1560
cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt   1620
tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgccccta   1680
ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc ccagtacatg accttatggg   1740
actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg gtgatgcggt   1800
tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc   1860
accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac tttccaaaat   1920
gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct   1980
atataagcag agctctctgg ctaactagag aacccactgc ttactggctt atcgaaatta   2040
atacgactca ctatagggag acccaagctg gctagcgttt aaacttaagc ttgccaccat   2100
gaagctactg tcttctatcg aacaagcatg cgatatttgc cgacttaaaa agctcaagtg   2160
ctccaaagaa aaaccgaagt gcgccaagtg tctgaagaac aactgggagt gtcgctactc   2220
tcccaaaacc aaaaggtctc cgctgactag ggcacatctg acagaagtgg aatcaaggct   2280
agaaagactg gaacagctat ttctactgat ttttcctcga gaagaccttg acatgatttt   2340
gaaaatggat tctttacagg atataaaagc attgttaaca ggattatttg tacaagataa   2400
tgtgaataaa gatgccgtca cagatagatt ggcttcagtg gagactgata tgcctctaac   2460
attgagacag ttgactgtat cgccgggtac cgagctcgga tccactagtc cagtgtggtg   2520
gaattctgca gatatccagc acagtggcgg ccgctcgagt ctagagggcc cgtttaaacc   2580
cgctgatcag cctcgactgt gccttctagt tgccagccat ctgttgtttg cccctccccc   2640
gtgccttcct tgaccctgga aggtgccact cccactgtcc tttcctaata aaatgaggaa   2700
attgcatcgc attgtctgag taggtgtcat tctattctgg ggggtggggt ggggcaggac   2760
agcaaggggg aggattggga agacaatagc aggcatgctg gggatgcggt gggctctatg   2820
gcttctgagg cggaaagaac cagctggggc tctagggggt atccccacgc gccctgtagc   2880
ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac acttgccagc   2940
gccctagcgc ccgctccttt cgctttcttc ccttcctttc tcgccacgtt cgccggcttt   3000
ccccgtcaag ctctaaatcg ggggctccct ttagggttcc gatttagtgc tttacggcac   3060
ctcgacccca aaaaacttga ttagggtgat ggttcacgta gtgggccatc gccctgatag   3120
acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact cttgttccaa   3180
actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg gattttgccg   3240
atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc gaattaattc   3300
tgtggaatgt gtgtcagtta gggtgtggaa agtccccagg ctccccagca ggcagaagta   3360
tgcaaagcat gcatctcaat tagtcagcaa ccaggtgtgg aaagtcccca ggctccccag   3420
caggcagaag tatgcaaagc atgcatctca attagtcagc aaccatagtc ccgcccctaa   3480
ctccgcccat cccgccccta actccgccca gttccgccca ttctccgccc catggctgac   3540
taattttttt tatttatgca gaggccgagg ccgcctctgc ctctgagcta ttccagaagt   3600
agtgaggagg cttttttgga ggcctaggct tttgcaaaaa gctcccggga gcttgtatat   3660
ccattttcgg atctgatcaa gagacaggat gaggatcgtt tcgcatgatt gaacaagatg   3720
gattgcacgc aggttctccg gccgcttggg tggagaggct attcggctat gactgggcac   3780
aacagacaat cggctgctct gatgccgccg tgttccggct gtcagcgcag gggcgcccgg   3840
ttcttttgt caagaccgac ctgtccggtg ccctgaatga actgcaggac gaggcagcgc   3900
ggctatcgtg gctggccacg acgggcgttc cttgcgcagc tgtgctcgac gttgtcactg   3960
aagcgggaag ggactggctg ctattgggcg aagtgccggg gcaggatctc ctgtcatctc   4020
accttgctcc tgccgagaaa gtatccatca tggctgatgc aatgcggcgg ctgcatacgc   4080
ttgatccggc tacctgccca ttcgaccacc aagcgaaaca tcgcatcgag cgagcacgta   4140
ctcggatgga agccggtctt gtcgatcagg atgatctgga cgaagagcat caggggctcg   4200
cgccagccga actgttcgcc aggctcaagg cgcgcatgcc cgacggcgag gatctcgtcg   4260
tgacccatgg cgatgcctgc ttgccgaata tcatggtgga aaatggccgc ttttctggat   4320
tcatcgactg tggccggctg ggtgtggcgg accgctatca ggacatagcg ttggctaccc   4380
gtgatattgc tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg ctttacggta   4440
tcgccgctcc cgattcgcag cgcatcgcct ctatcgcct tcttgacgag ttcttctgag   4500
cgggactctg gggttcgaaa tgaccgacca agcgacgccc aacctgccat cacgagattt   4560
cgattccacc gccgccttct atgaaaggtt gggcttcgga atcgttttcc gggacgccgg   4620
ctggatgatc ctccagcgcg gggatctcat gctggagttc ttcgcccacc ccaacttgtt   4680
tattgcagct tataatggtt acaaataaag caatagcatc acaaatttca caaataaagc   4740
atttttttca ctgcattcta gttgtggttt gtccaaactc atcaatgtat cttatcatgt   4800
ctgtataccg tcgacctcta gctagagctt ggcgtaatca tggtcatagc tgtttcctgt   4860
gtgaaattgt tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa   4920
```

```
agcctggggt gcctaatgag tgagctaact cacattaatt gcgttgcgct cactgcccgc     4980
tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag     5040
aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt     5100
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga     5160
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg     5220
taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacaa     5280
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt     5340
tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct     5400
gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct     5460
cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc     5520
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt     5580
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc     5640
tacagagttc ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat     5700
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa     5760
acaaaccacc gctggtagcg gttttttttgt ttgcaagcag cagattacgc gcagaaaaa     5820
a                                                                     5821
```

<210> 41
<211> 10075
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
     DNA molecule sequence

<220>
<221> modified_base
<222> (9040)..(9040)
<223> unsequenced nucleotide

<400> 41
```
ctggtagcgg tggtttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc       60
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt      120
aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa      180
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat      240
gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct      300
gactccccgt cgtgtagata actacgatac gggagggctt accatctggc cccagtgctg      360
caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata aaccagccag      420
ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta      480
attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg      540
ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg      600
gttcccaacg atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa gcggttagct      660
ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta      720
tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg      780
gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc      840
cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg      900
gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga      960
tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg     1020
ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat     1080
gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc     1140
tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca     1200
catttccccg aaaagtgcca cctaaattgt aagcgttaat attttgttaa aattcgcgtt     1260
aaatttttgt taaatcagct catttttttaa ccaataggcc gaaatcggca aaatccctta     1320
taaatcaaaa gaatagaccg agatagggtt gagtgttgtt ccagtttgga acaagagtcc     1380
actattaaag aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg     1440
cccactacgt gaaccatcac cctaatcaag ttttttgggg tcgaggtgcc gtaaagcact     1500
```

```
aaatcggaac cctaaaggga gcccccgatt tagagcttga cggggaaagc cggcgaacgt    1560
ggcgagaaag gaagggaaga aagcgaaagg agcgggcgct agggcgctgg caagtgtagc    1620
ggtcacgctg cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac agggcgcgtc    1680
ccattcgcca ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg cctcttcgct    1740
attacgccag ctggcgaaag ggggatgtgc tgcaaggcga ttaagttggg taacgccagg    1800
gttttcccag tcacgacgtt gtaaaacgac ggccagtgag cgcgcgtaat acgactcact    1860
atagggcgaa ttggagctcc accgcggtgc ggccgggcca tgcaggccac gacatgataa    1920
gatacattga tgagtttgga caaaccacaa ctagaatgca gtgaaaaaaa tgctttattt    1980
gtgaaatttg tgatgctatt gctttatttg taaccattat aagctgcaat aaacaagttc    2040
ttgtacagct cgtccatgcc gagagtgatc ccggcggcgg tcacgaactc cagcaggacc    2100
atgtgatcgc gcttctcgtt ggggtctttg ctcagggcgg actggtagct caggtagtgg    2160
ttgtcgggca gcagcacggg gccgtcgccg atggggtgt tctgctggta gtggtcggcg    2220
agctgcacgc tgccgtcctc gatgttgtgg cggatcttga agttccctt gatgccgttc    2280
ttctgcttgt cggccatgat atagacgttg tggctgttgt agttgtactc cagcttgtgc    2340
cccaggatgt tgccgtcctc cttgaagtcg atgcccttca gctcgatgcg gttcaccagg    2400
gtgtcgccct cgaacttcac ctcggcgcgg gtcttgtagt tgccgtcgtc cttgaagaag    2460
atggtgcgct cctggacgta gccttcgggc atggcggact tgaagaagtc gtgctgcttc    2520
atgtggtcgg ggtagcgggc gaagcactgc aggccgtagc cgaaggtggt cacgagggtg    2580
ggccagggca cgggcagctt gccggtggtg cagatgaact tcagggtcag cttgccgtag    2640
gtggcatcgc cctcgccctc gccggacacg ctgaacttgt ggccgtttac gtcgccgtcc    2700
agctcgacca ggatgggcac caccccggtg aacagctcct cgcccttgct caccatggtg    2760
gcgttagctt gatttgacag tggctggggg ttgcgccgcc gggtttata ggaagccaca    2820
gcggccactc gagccataaa aggcaacttt aggaacggcg ggggtgatt ggattcgagt    2880
cgtttattca ccggccttgc cgcacagtgc agcatttttt tacccctct cccctccttt    2940
tgcggggaaa aaaaaaaaa aaaaaaaaa aaggagaaga gaaaaaagc gagcgagaga    3000
gaaagcgaga ttgaggaaga ggatgaagag tttggcgatg ggtgctggtt ccgtaggccc    3060
agatggacaa gaatagcccc cgcccttgcg gacagtatcc cattcagtga ctcagatcag    3120
atcaagcggc cgccagtgtg atggatatct gcagaattcc agcacactgg cggccgttac    3180
tagtggatcc ctctggagct gcctagtgtg ggtgtctgct ggaaacttct aaaatgtctc    3240
ggcctattta attggagaat tcactgaaga attaaaggac tcagtaaaga agaggtttgg    3300
tttgttcatg aaagtaagta gcttgtgaaa tgctttattt tcctgtgact aaaattccac    3360
cttgtaattc ccaataagga ttgcttcaag ctatttaata caaattaaac ctgtttcacc    3420
atgctcatag gaaaatgtca aagtatcttc attcaccaac tgagggctgg ctcttctttg    3480
aaggtagctc tgagatttgt tttgtaatgg agacaaagac ttatttaact gttggggagg    3540
gggaaggagg acttaggcca ggcaggcggg gatgtggggg acgaaccatg aggggcaatt    3600
tggctgcctc tgaaaaggtc cagagctacc ttaaaagaca ggctttctat aacaagtggt    3660
cctttctttt cctttctttt ttttttgag acagagtctc actctgttgc ccaagctgga    3720
gtgcagtggt gcgatctcag ctcactgcaa cctccagctc ccaggttcaa gcaattctcc    3780
tgcctcagcc tcccaagtag ctgggattac aggcaaccgc taccatgccc agctaatttt    3840
ttgtattttt aatagagatg gggtttcacc atgttggcca gggtggtctc aaactcctga    3900
cctcaagtga tccacccacc tcagcctccc aaagtgctgg gattacagac atgagccacc    3960
gcgcccagac acagctggct ttgtcgccca ggctggaagt gcagtggcac gatctcggct    4020
cactgcaagc tctgcctcct gggttcacgc cattctcctg cctcagcttc caagtagct    4080
gagactacag gcgcccgcca ccacgcccgg ctaattttt gtattttag tggagatggt    4140
gtttcaccat gttagccagg atggtcttga tctcctgacc tcgtgatccg cctgtcttgg    4200
cctcccaaag tgctgggatg acaggcgtga gccaccacgc ctggcttaca agtggccctt    4260
tctaagcccc aagatgtcac tgtcttttca ctgaaggcgg aaaagaatg taggcatgtt    4320
tgtctgagag ggagggagta gtttttaggag aaaggaccaa gagcacattt tgagtgggaa    4380
gtaagtggtt tccttttcta gccttatgat atggtttggc tctgtgtccc cacccaaatc    4440
tcatggtgaa ctgtaatccc cagtggtgga ggtgaggcct ggcgggaggg actggatcat    4500
gggggtgatc cttcatgaat ggtttagcac catcctcttg gtgctactct agtgatagcg    4560
ttctcacaag atctgatggg tgaaaagtgt gtgtctcttg ttttgctcc tgccatgcaa    4620
gacgcctact cccacttcac cttctgctat gagtcaaatt ccctgaggcc tccccagaag    4680
cagaagcttc ctgtacagcc ttgtgggcca cgagccaatt aaacctcttt tctttataaa    4740
ttaccaggtc tcagggattt ctttatagca gactaataca cctcaaagta gggatttatt    4800
ttattattct attttttttt tgagataggg tctcgatctg tcacccaggc tggagtgcag    4860
tggcatgatc tcagcttgct gtaacctctg cctcccgggt tgaagcaatt ctcatgccta    4920
agccacccaa gtagctggaa ttacaggtgt gtgccaccct gctcagctaa ttttttgtatc    4980
```

```
tttagtagag ttggggtttc tccatgttgg ccaggctggt atcaaactcc tggactcaag   5040
caatctgccc accttggcct ccccagagtg ccaggattac aggcatgacc catcttgcct   5100
ggtcagtttt ttttttttta aatccagtta atgctgggtg ccatgactca cacctgtaat   5160
gccagtgctt tgggaggccg aggtgggaag atcacttgag gccaggagtt tgagaccagc   5220
ctgggaaaca tctcatctct accaaaaaac aaaaacaaaa caaaacaaca acaacatcaa   5280
acatagccag gtgtggtggt gcacacctgt agtcccagct ccttgggagg ctgaggcagg   5340
aggattgctt gagcccagga gttcaaggct gcagtgaact atgattgcac cactgcattc   5400
cagcctgggc aacagagcaa gaacctatct ctatttaaga aaaaatgctc ctaagaaatc   5460
atcttagtac ttggtaagct attaaggttt tttttaagtg tattaattac atagccagta   5520
taataactta gttgctaaga gctttgaggt ggctccaagc aagaagaaag gagttgttgt   5580
taagctacag ctcagagtct cagaatgaac tgtttccatc ttcgctgtcc tttgcttaac   5640
tgtgggataa atataatttt ctcaacattc aacacatgtg tccttaatcg tcctctacga   5700
cactggttcc atgttgccaa gagaacccag aacagccgca aacacctact ggattgttac   5760
tggcctttcc tgagaggatg acccgggctt tgaccttgtt catgttgaag ttcttcaagt   5820
aggcctcgta gattctcttg ccagagatt tgagatctgc agtttcagaa tcttctatgt   5880
catgttcaca ggtaagaatt tctgctttca gttttgcttt ctcagatctt ggcattcgtc   5940
cggtacccgg cgatacagtc aactgtcttt gacctttgtt actactctct tccgatgatg   6000
atgtcgcact tattctatgc tgtctcaatg ttagaggcat atcagtctcc actgaagcca   6060
atctatctgt gacggcatct ttattacat tatcttgtac aaataatcct gttaacaatg   6120
cttttatatc ctgtaaagaa tccattttca aaatcatgtc aaggtcttct cgaggaaaaa   6180
tcagtagaaa tagctgttcc agtctttcta gccttgattc cacttctgtc agatgtgccc   6240
tagtcagcgg agaccttttg gttttgggag agtagcgaca ctcccagttg ttcttcagac   6300
acttggcgca cttcggtttt tctttggagc acttgagctt tttaagtcgg caaatatcgc   6360
atgcttgttc gatagaagac agtagcttca tggtggcaag cttaagttta aacgctagaa   6420
ctagtggatc cggataagcc agtaagcagt gggttctcta gttagccaga gagctctgct   6480
tatatagacc tcccaccgta cacgcctacc gcccatttgc gtcaatgggg cggagttgtt   6540
acgacatttt ggaaagtccc gttgattttg gtgccaaaac aaactcccat tgacgtcaat   6600
ggggtggaga cttggaaatc cccgtgagtc aaaccgctat ccacgcccat tgatgtactg   6660
ccaaaaccgc atcaccatgg taatagcgat gactaatacg tagatgtact gccaagtagg   6720
aaagtcccat aaggtcatgt actgggcata atgccaggcg ggccatttac cgtcattgac   6780
gtcaataggg ggcgtacttg gcatatgata cacttgatgt actgccaagt gggcagttta   6840
ccgtaaatac tccacccatt gacgtcaatg gaaagtccct attggcgtta ctatgggaac   6900
atacgtcatt attgacgtca atgggcgggg gtcgttgggc ggtcagccag gcgggccatt   6960
taccgtaagt tatgtaacgc ggaactccat atatgggcta tgaactaatg accccgtaat   7020
tgattactat taataactag tcaataatca atgtcaaccc cgggctgcag gaattctacc   7080
gggtagggga ggcgcttttc ccaaggcagt ctggagcatg cgctttagca gccccgctgg   7140
cacttggcgc atcacaagtg gcctctggcc tcgcacacat tccacatcca ccggtagcgc   7200
caaccggctc ccttctttgg tggccccttc gcgccacctt ctactcctcc cctagtcagg   7260
aagttccccc ccgccccgca gctcgcgtcg tgcaggacgt gacaaatgga agtagcacgt   7320
ctcactagtc tcgtgcagat ggacaagcac cgctgagcaa tggaagcggg taggcctttg   7380
gggcagcggc aatagcagc ttggctcctt cgctttctgg gctcagaggc tgggaagggg   7440
tgggtccggg ggcgggctca ggggcgggct caggggcggg cgggcgcga aggtcctccg   7500
gacccggcat tctgcacgct tcaaaagcgc acgtctgccg cgctgttctc ctcttcctca   7560
tctccgggcc ttcgacctgc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt   7620
ttctgatcga aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat   7680
ctcgtgcttt cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg   7740
ccgatggttt ctacaaagat cgttatgttt atcggcactt tgcatcggcc gcgctcccga   7800
ttccggaagt gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc   7860
gtgcacaggg tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc   7920
cggtcgcgga ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg   7980
gcccattcgg accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga   8040
ttgctgatcc ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg   8100
tcgcgcaggc tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc   8160
tcgtgcacgc ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg   8220
tcattgactg gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct   8280
tctggaggcc gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc   8340
cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac   8400
tctatcagag cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg   8460
```

```
acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg    8520
cggccgtctg gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca    8580
gcactcgtcc gagggcaaag gaatgcctga gaaaggaagt gagctgtaaa ggctgagctc    8640
tctctctgac gtatgtagcc tctggttagc ttcgtcactc actgttcttg actcagcatg    8700
gcaatctgat gaaatcccag ctgtaagtct gcataaattg atgatctatt aaacaataaa    8760
gatgtccact aaaatggaag tttttactg tcatactttg ttaagaaggg tgagaacaga    8820
gtacctacat tttgaatgga aggattggag ctacgggggt gggggtgggg gtgggattag    8880
ataaatgcct gctctttact gaaggctctt tactattgct ttatgataat gtttcatagt    8940
tggatatcat aatttaaaca agcaaaacca aattaagggc cagctcattc ctccactcat    9000
gatctataga tctatagatc tctcgtggga tcattgtttn tctgatccac tggaagctta    9060
tcgataccgt cgacctcgag gggggggcccg gtacccagct tttgttccct ttagtgaggg    9120
ttaattgcgc gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg    9180
ctcacaattc cacacaacat acgagccgga agcataaagt gtaaagcctg gggtgcctaa    9240
tgagtgagct aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac    9300
ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt    9360
gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg ctgcggcga    9420
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca    9480
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg    9540
ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg acgctcaagt    9600
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc    9660
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct    9720
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc    9780
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta    9840
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca    9900
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag    9960
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag   10020
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccg         10075
```

<210> 42
<211> 8971
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
      DNA vector sequence

<400> 42
```
gtaggtgtca ttctattctg gggggtgggg tggggcagga cagcaagggg gaggattggg      60
aagacaatag caggcatgct ggggatgcgg tgggctctat ggcttctgag gcggaaagaa     120
ccagctgggg ctctaggggg tatccccacg cgccctgtag cggcgcatta agcgcggcgg     180
gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt     240
tcgctttctt cccttccttt ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc     300
gggggctccc tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg     360
attagggtga tggttcacgt agtgggccat cgccctgata gacggttttt cgccctttga     420
cgttggagtc cacgttcttt aatagtggac tcttgttcca aactggaaca acactcaacc     480
ctatctcggt ctattctttt gatttataag ggattttgcc gatttcggcc tattggttaa     540
aaaatgagct gatttaacaa aaatttaacg cgaattaatt ctgtggaatg tgtgtcagtt     600
agggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca tgcatctcaa     660
ttagtcagca accaggtgtg gaaagtcccc aggctcccca gcaggcagaa gtatgcaaag     720
catgcatctc aattagtcag caaccatagt cccgccccta actccgccca tcccgccccт     780
aactccgccc agttccgccc attctccgcc ccatggctga ctaattttt ttatttatgc     840
agaggccgag gccgcctctg cctctgagct attccagaag tagtgaggag gcttttttgg     900
aggcctaggc ttttgcaaaa agctcccggg agcttgtata tccattttcg gatctgatca     960
gcacgtgatg aaaaagcctg aactcaccgc gacgtctgtc gagaagtttc tgatcgaaaa    1020
gttcgacagc gtctccgacc tgatgcagct ctcggagggc gaagaatctc gtgctttcag    1080
cttcgatgta ggagggcgtg gatatgtcct gcgggtaaat agctgcgccg atggtttcta    1140
```

```
caaagatcgt tatgtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgct   1200
tgacattggg gaattcagcg agagcctgac ctattgcatc tcccgccgtg cacagggtgt   1260
cacgttgcaa gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc   1320
catggatgcg atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc   1380
gcaaggaatc ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca   1440
tgtgtatcac tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct   1500
cgatgagctg atgctttggg ccgaggactg ccccgaagtc cggcacctcg tgcacgcgga   1560
tttcggctcc aacaatgtcc tgacggacaa tggccgcata acagcggtca ttgactggag   1620
cgaggcgatg ttcgggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg   1680
gttggcttgt atggagcagc agacgcgcta cttcgagcgg aggcatccgg agcttgcagg   1740
atcgccgcgg ctccgggcgt atatgctccg cattggtctt gaccaactct atcagagctt   1800
ggttgacggc aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgtccg   1860
atccggagcc gggactgtcg ggcgtacaca aatcgcccgc agaagcgcgg ccgtctggac   1920
cgatggctgt gtagaagtac tcgccgatag tggaaaccga cgccccagca ctcgtccgag   1980
ggcaaaggaa tagcacgtgc tacgagattt cgattccacc gccgccttct atgaaaggtt   2040
gggcttcgga atcgtttttcc gggacgccgg ctggatgatc ctccagcgcg gggatctcat   2100
gctggagttc ttcgcccacc ccaacttgtt tattgcagct tataatggtt acaaataaag   2160
caatagcatc acaaatttca caaataaagc attttttttca ctgcattcta gttgtggttt   2220
gtccaaactc atcaatgtat cttatcatgt ctgtataccg tcgacctcta gctagagctt   2280
ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca   2340
caacatacga gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact   2400
cacattaatt gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct   2460
gcattaatga atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc   2520
ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca   2580
ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg   2640
agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca   2700
taggctccgc ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa   2760
cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc   2820
tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc   2880
gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct   2940
gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg   3000
tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag   3060
gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta   3120
cggctacact agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg   3180
aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtttttttgt   3240
ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc   3300
tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt   3360
atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta   3420
aagtatatat gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat   3480
ctcagcgatc tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac   3540
tacgatacgg gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg   3600
ctcaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag   3660
tggtcctgca actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt   3720
aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt   3780
gtcacgctcg tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt   3840
tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt   3900
cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct   3960
tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt   4020
ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac   4080
cgcgccacat agcagaactt taaaagtgct catcattgga aaacgttctt cggggcgaaa   4140
actctcaagg atcttaccgc tgttgagatc cagttcgatg taacccactc gtgcacccaa   4200
ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca   4260
aaatgccgca aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct   4320
ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga   4380
atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc   4440
tgacgtcgac ggatcgggag atctcccgat cccctatggt gcactctcag tacaatctgc   4500
tctgatgccg catagttaag ccagtatctg ctccctgctt gtgtgttgga ggtcgctgag   4560
tagtgcgcga gcaaaattta agctacaaca aggcaaggct tgaccgacaa ttgcatgaag   4620
```

```
aatctgctta gggttaggcg ttttgcgctg cttcgcgatg tacgggccag atatacgcgt   4680
tgacattgat tattgactag ttattaatag taatcaatta cggggtcatt agttcatagc   4740
ccatatatgg agttccgcgt tacataactt acggtaaatg gcccgcctgg ctgaccgccc   4800
aacgaccccc gcccattgac gtcaataatg acgtatgttc ccatagtaac gccaataggg   4860
actttccatt gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt ggcagtacat   4920
caagtgtatc atatgccaag tacgcccccт attgacgtca atgacggtaa atggcccgcc   4980
tggcattatg cccagtacat gaccttatgg gactttccta cttggcagta catctacgta   5040
ttagtcatcg ctattaccat ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag   5100
cggtttgact cacggggatt tccaagtctc caccccattg acgtcaatgg gagtttgttt   5160
tggcaccaaa atcaacggga ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa   5220
atgggcggta ggcgtgtacg gtgggaggtc tatataagca gagctctctg gctaactaga   5280
gaacccactg cttactggct tatcgaaatt aatacgactc actataggga gacccaagct   5340
ggctagcgtt taaacttaag cttgccacca tgaagctact gtcttctatc gaacaagcat   5400
gcgatatttg ccgacttaaa aagctcaagt gctccaaaga aaaaccgaag tgcgccaagt   5460
gtctgaagaa caactgggag tgtcgctact ctcccaaaac caaaaggtct ccgctgacta   5520
gggcacatct gacagaagtg gaatcaaggc tagaaagact ggaacagcta tttctactga   5580
tttttcctcg agaagacctt gacatgattt tgaaaatgga ttctttacag gatataaaag   5640
cattgttaac aggattattt gtacaagata atgtgaataa agatgccgtc acagatagat   5700
tggcttcagt ggagactgat atgcctctaa cattgagaca gcatagaata agtgcgacat   5760
catcatcgga agagagtagt aacaaaggtc aaagacagtt gactgtatcg ccgggtaccg   5820
cagtggggat gtctcataat ggtaagtaaa cagtcatcac catatacttt attattctca   5880
ttatagctgc cagaccagtg gacactaaag ccattgccaa aaatgtgtac agtttttcca   5940
ccaaatgcca gaatttagaa tattgcatgg cgataaaaca tttctctttt aggtcagtgt   6000
tttttaaagtt ttattataga acctttctct ctgtggttgg gcatctgcca tgaggagaaa   6060
agagacttga aaaatctggg ggattatggg aaaaacttta gtgttagaac aaatataatg   6120
atcatcatga ttaattttta tattgcctca tgcattaaaa acagaattaa aatttacttt   6180
tgtgtatcag tagtgtcatt ctgctttaaa tgtaagaaaa tcctgaagat gtctgggtct   6240
gaaactttaa atggcatatg agaatttaga ttcatttaat ttttttttttt ttttttttttt   6300
tgctgtttat tatgcaaaga cgagtattcg tttccttcac ttcaaggaac gtccatgagt   6360
tgtggaaagg gtaaaacaaa ttattttat agttctgcaa aacttcaaag ccatagaaac   6420
ataagagaat aaaactttct acttttgaaa cctgtgtcta aagtcaagct tatgactttg   6480
ggtgtcaact taagggaaat gatgtagaag acctgagaaa agacccatgt tttcatgatc   6540
tgtggaacta actcagtaag gtcctggaag acaaacgtgt ttgaaaaaat aattcttggt   6600
cactatttgg ttctgattgg ttgcctggct gacaaaataa attttaggtg taacttgata   6660
gtattatatg tgtctagtat taatcagtta ttttcaatgc actcaaggaa taaaaaatgc   6720
caacatgatc aaattaaatt tgaagtttca tttcatcagc tatcctattt tttctttctt   6780
ttgtttttcat ctctctcaaa actgatctag taagaaggca gtcttcatct atatgtactc   6840
taaagcatct gagaaaatta aatttttgct ctatcatgca gatacttggt tttcacagtg   6900
gaaaccatca agatatttga gagtcactca ttgattaata acctaccctt taaaaatagt   6960
tcacaaaatt ctatctgtat gattactaat gggtcagttc ctctaaaggc aagaatgtat   7020
tttgaataga gctatcagca gtactgaaca ggatggaaat ttggtccaac tttcctcaac   7080
aattgactac atctgctgta tacaactgta gtcatggatt gcaaaataat ttgctggttc   7140
agttccatca gtttcctttt tatttcagac ctcataaaac ataagggcaa gcttccaagt   7200
caagggttgt attttaggtg atgcaatcaa gttgcaataa acttatttca tcttcatgaa   7260
gaaatcattg actttgaggt cattatcact gagacagcct aggaaagcat gccagttccg   7320
cttttaacac taattttttgc ctattaaaaa attatcactt cacacattca tgtaaaatat   7380
aattcctgtt acttacaggc agtgcataca ttgatgtaaa atatatttcc tgttgcttac   7440
aggcattgca tacgtttact taaaatcaat ctgcaattgc ttattgttta taggcaaata   7500
ttatttaaac cccaacctat gtagtatgga tggtataatt ttctcctccc ataaatgtca   7560
aagaaagcta tcataaactt gataggtcat attttgactt catatgacaa tttattttat   7620
ttctagagag ggtcaatctt gatagggagt gtttgttcag ttctagcaaa agcaagaaaa   7680
tatatgaatt aagtgtcatg acatcagaga cttaggcagt tcaacactta aaaaattgca   7740
aacataacta agaaaacctc agaggttttt atgggagttt gtttgctgca cagaaaataa   7800
acatcagact aacaacaagg ttttgattag ttttgcctgt tagctgtatg aagaaaacat   7860
ttctaaatac aaagttttct aggaagttga gtcactattt aaaatccatt gctttgacca   7920
atgtgaagcc aacagggttt tcattcatat agaatgcctt tttatcaaaa gaggctctga   7980
gtgttatacc aaaatcagca cccagctgtt ttcagtatgc gcaagaaaac atgaacgtcc   8040
cccaaagcaa atgtactcca ttccgtgtta accttatttc tgctccctgg tctctgatta   8100
```

```
gtctggttga tgaatgtagc ttcccctttg gccatgcact cacagctttg aaggaggctt    8160
ctgaaggaac tggtagggaa ccctaggggc tcttgctaat gttgttggat agtcactgaa    8220
atcagtgatc accatgtcct ttcactttga aaattctatg attctgtata atcaaacaac    8280
ctaactcact ggattttca cttttcaacc agctattccg tgtcttacag tgctctggtc    8340
ttcacagata ggtaggcacc tggcatgtca gttgcctgaa tttgaaagtt ttcacctgta    8400
tgttttggta cgataaaaat aaaaatgtaa tttatatatc tgaatcaggt ctgtatgtta    8460
tgatcaattg ctcagcaatt tcgggcagtt ggtttgatgg ttatgtagta atgtagcctg    8520
agagcagaaa tacagagcct ctgggctaga gaaagtataa atggcatcct aggctatgta    8580
gggttacagc tcttcagaag gaactttcat tttcattgtg acacatcgtc tacatgttgt    8640
agaagaacat agtttcagaa ttcttccagt tagaaacata gtttctcaaa tattcacttt    8700
cagcattggg tagaaaaagt acccaggtga tggactgatt gaggatccac tagtaacggc    8760
cgccagtgtg ctggaattct gcagatatcc atcacactgg cggccgctcg agtctagagg    8820
gcccgtttaa acccgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt    8880
ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta    8940
ataaaatgag gaaattgcat cgcattgtct g                                   8971
```

```
<210> 43
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 43
cgggtccccg gcgatacagt caactgtct                                        29


<210> 44
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 44
taaagcttgc caccatgaag ctactgtctt                                       30


<210> 45
<211> 10514
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
      DNA vector sequence

<400> 45
ctaaattgta agcgttaata ttttgttaaa attcgcgtta aattttttgtt aaatcagctc    60
atttttttaac caataggccg aaatcggcaa atcccttat aaatcaaaag aatagaccga   120
gatagggttg agtgttgttc cagtttggaa caagagtcca ctattaaaga acgtggactc   180
caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc   240
ctaatcaagt ttttgggggt cgaggtgccg taaagcacta atcggaacc ctaaagggag    300
cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa   360
agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac   420
```

```
cacacccgcc gcgcttaatg cgccgctaca gggcgcgtcc cattcgccat tcaggctgcg    480
caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg    540
gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt cacgacgttg    600
taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaat tggagctcca    660
ccgcggtgcg gccgggccat gcaggccaga catgataaga tacattgatg agtttggaca    720
aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc    780
tttatttgta accattataa gctgcaataa acaagttgtc gagttacttg tacagctcgt    840
ccatgccgag agtgatcccg gcggcggtca cgaactccag caggaccatg tgatcgcgct    900
tctcgttggg gtctttgctc agggcggact ggtagctcag gtagtggttg tcgggcagca    960
gcacggggcc gtcgccgatg ggggtgttct gctggtagtg gtcggcgagc tgcacgctgc   1020
cgtcctcgat gttgtggcgg atcttgaagt tcaccttgat gccgttcttc tgcttgtcgg   1080
ccatgatata gacgttgtgg ctgttgtagt tgtactccag cttgtgcccc aggatgttgc   1140
cgtcctcctt gaagtcgatg cccttcagct cgatgcggtt caccagggtg tcgccctcga   1200
acttcacctc ggcgcgggtc ttgtagttgc cgtcgtcctt gaagaagatg gtgcgctcct   1260
ggacgtagcc ttcgggcatg gcggacttga agaagtcgtg ctgcttcatg tggtcggggt   1320
agcgggcgaa gcactgcagg ccgtagccga aggtggtcac gagggtgggc cagggcacgg   1380
gcagcttgcc ggtggtgcag atgaacttca gggtcagctt gccgtaggtg gcatcgccct   1440
cgccctcgcc ggacacgctg aacttgtggc cgtttacgtc gccgtccagc tcgaccagga   1500
tgggcaccac cccggtgaac agctcctcgc ccttgctcac catggtggcg ttagcttgat   1560
ttgacagtgg ctgggggttg cgccgccggg ttttatagga agccacagcg gccactcgag   1620
ccataaaagg caactttagg aacggcgggg ggtgattgga ttcgagtcgt ttattcaccg   1680
gccttgccgc acagtgcagc atttttttac cccctctccc ctccttttgc gggggaaaaa   1740
aaaaaaaaaa aaaaaaaaag gagaagagaa aaaaagcgag cgagagagaa agcgagattg   1800
aggaagagga tgaagagttt ggcgatgggt gctggttccg taggcccaga tggacaagaa   1860
tagcccccgc ccttgcggac agtatcccat tcagtgactc agatcagtca agcggccgcg   1920
tctaacctag gagcccacca cagacactat actataacca cactgaattt tctctcctat   1980
ctcaaataat cctattcttg ttcatgcctc agtgccttat aacaagcaga tttctcggcc   2040
agacatggtg gcttgcgcct gtaatcccag cactttggcc aagtcgggca gctcacttga   2100
ggtcaggagt ttgaattcag cctgtccaat atggtgaaac cccatctcta ctaaaaatac   2160
aaaaattagc tgggcatggt ggtgtgtgcc tgtagtccca gctactcagg agactgaggc   2220
aggagaattg cttgaactcg ggaggcagag gttgcagtga gccaagatca tgccattgca   2280
ctccagcctg ggcaacagag cgagactccg tctcaaacaa aacaaaacaa aaacaaaaaa   2340
acaagcagat ctctcaactt ggaatgctgc tgcttctcca tgccaactgg aaaacttctt   2400
tttctttttag actcaactca aatatgatcc tccctgaagt ctctccatct ttcctagaca   2460
caattagcaa ccagttcctc tgtgttcaaa cagcacttcg tgtgtatgtg tatgtatgtg   2520
tatttatcat tattattatc acctttgatc actctccctg aaggtaatag ccaggcactt   2580
agtatagagt aggtatcaac aattgcttat tgaaggaata aatgggtgaa tgacaacttt   2640
gcaaagtttc aatttttaat ctattgattt gaaagaatta ttatacatta ttttaaatat   2700
cctttctagt tctaatatct tatgtctata gtagttcgac agttacttta aatcagatat   2760
taaatgattt tcattttaca aagaacagac ctgcctttct caatgtcaca agactattgt   2820
agagaatagc agtgaaataa gtacatgtga aaatacttgg caaagtaata gtgaaaatag   2880
agctgatctc acaggaatgt tttgaaaatt aaaacaccgt atgtaaaata ggtaatctaa   2940
tctgtgggca catatttaaa ctctcaataa atattaactt tgattatttta aaaaatacct   3000
ggtagctttt ttgtgaattc tacaagaacc tgtacatgat tggttgccat ttccgtcaaa   3060
atgagaaaat tttcttctgc actgaattct tcttttaact aaattttaag aaaaaaacca   3120
atgattaaaa agtagataag aaaatgataa tataaaacca ttagactaaa gaaaaactag   3180
gtttagatga taaaatttga aaacccatct tactatttct cagatctttg gagggcaatt   3240
accattatca agataggtgg gagacttcat actcttgaat tgtttaatgt tgctgagagt   3300
ttaaaagtct aacttgaatt ctcagctcta atcatttcaa agtaacatat tcacatataa   3360
attatacctc caaggcaata aaaactccag aaaatattaa ctttaaaccc acatattttc   3420
agatattata catacaattt tatttgttat ttcctgaggc atcctctgtt tgttatatga   3480
atccataata aaatgtagaa gagtctgttg atctggggtg agttcagttt tctcctacac   3540
tggcaagaaa ataaaaggtg ttagggaaga tgtgtataaa gccttctttc tatatgtgca   3600
ttcatctgga gggacctaat agcagaaatc ataaaaaaaa aaaggcataa atacagcagg   3660
aactatgagt tctaagagag tgtataattc ctttctctgg gtcctcaggc aggccatatc   3720
acatttcact gaattttcct atctataaaa tatgaataac aaaatattcc acttgtgtat   3780
ttcttaggaa taaatggtag tctggatatg aagactattg agcccattca tattgaaggt   3840
cactgaaagt tcaagcttta attataacta atgaaaggtt ttcatatagt atgactacta   3900
```

```
tgaatactat attcatatac tattctacat gaattccacc tggtgaagtc atctgagagt    3960
attcatgtgg aggagggatc aacaaactgt ggcccatgga ccaaatccat tggatttat     4020
aaagttttat tggaacacag tcacactcat ttgtttatgt attgtctctg ctactttgc     4080
actaaaatgg cagagttgag tagctgttgt agataccatg tggccaaagc cctttaaaga    4140
aaaaaagttt gtcaacccct aaaatagaga gtcaacaatc tggacaattc agaccccaaa    4200
catctctttc tccctctaca gcatcttttg aagagtgtga ataattatgt caactctttg    4260
ttttcttata ccaaagtgtt actgatgaag cactatggca aaataatgct ttcatataaa    4320
atactttttc ttaaataaat acaaaaatac gtatttgtaa aactaaatgt ggataaagtt    4380
ttctttgtca ataatgtaat aaatattgag aagggtgtgg tttggacaat catagcttac    4440
tggtttttca gctaatggca aaataaaatt ccatcaaatg tcctgtgcat tatcacaatg    4500
tgtgtgtcta aacaccttgg atgctgttcc cctattttgt cctaagttgc tcaatggatt    4560
tcttgagaaa tttgtaaccc cagtgggtag gactttctc tattcatcca catacgtgta     4620
ttcaccctgt tttactcatt tataaggatt aatccacagc ctctattgag accatgttct    4680
tcccagtttg agactgtctt ctcaaaaaga gggaaaaaat ggtcatcaat actcagttgg    4740
aaaaaatcca ttttatacac caaatcattt ctgctggcaa acactaggta gtaggactgg    4800
aatggagtca gagagaagga aaactttcac tttttagttt atgcattgct actatttgaa    4860
tgtttacatc aaagcatgta ttaggaatcc attttaatga acaattggac tagaaattca    4920
gctgtgcctg tggctttaat ttgaaagcct cctggccagc tcactatgta cctagaaact    4980
cagttcctgc cagtcttggc agacaccatt gcatattatt accctgcatg actttgttgt    5040
cgaggtcact tgtcgcaagt cacgaccttc actgtcttct ctagacggcg atacagtcaa    5100
ctgtctttga cctttgttac tactctcttc cgatgatgat gtcgcactta ttctatgctg    5160
tctcaatgtt agaggcatat cagtctccac tgaagccaat ctatctgtga cggcatcttt    5220
attcacatta tcttgtacaa ataatcctgt taacaatgct tttatatcct gtaaagaatc    5280
cattttcaaa atcatgtcaa ggtcttctcg aggaaaaatc agtagaaata gctgttccag    5340
tctttctagc cttgattcca cttctgtcag atgtgcccta gtcagcggag accttttggt    5400
tttgggagag tagcgacact cccagttgtt cttcagacac ttggcgcact tcggttttc     5460
tttggagcac ttgagctttt taagtcggca aatatcgcat gcttgttcga tagaagacag    5520
tagcttcatg gtggcggatc ccgataagcc agtaagcagt gggttctcta gttagccaga    5580
gagctctgct tatatagacc tcccaccgta cacgcctacc gcccatttgc gtcaatgggg    5640
cggagttgtt acgacatttt ggaaagtccc gttgattttg gtgccaaaac aaactcccat    5700
tgacgtcaat ggggtggaga cttggaaatc cccgtgagtc aaaccgctat ccacgcccat    5760
tgatgtactg ccaaaaccgc atcaccatgg taatagcgat gactaatacg tagatgtact    5820
gccaagtagg aaagtcccat aaggtcatgt actgggcata atgccaggcg ggccatttac    5880
cgtcattgac gtcaataggg ggcgtacttg gcatatgata cacttgatgt actgccaagt    5940
gggcagttta ccgtaaatac tccacccatt gacgtcaatg gaaagtccct attggcgtta    6000
ctatgggaac atacgtcatt attgacgtca atgggcgggg gtcgttgggc ggtcagccag    6060
gcgggccatt taccgtaagt tatgtaacgc ggaactccat atatgggcta tgaactaatg    6120
accccgtaat tgattactat taataactag tcaataatca atgtcaacgc gtatatctgg    6180
cccgtacatc gcgaagcagc gcaaaacgcc taaccctacc cgggctgcag gaattctacc    6240
gggtagggga ggcgcttttc ccaaggcagt ctggagcatg cgctttagca gccccgctgg    6300
cacttggcgc atcacaagtg gcctctggcc tcgcacacat tccacatcca ccggtagcgc    6360
caaccggctc ccttctttgg tggccccttc gcgccacctt ctactcctcc cctagtcagg    6420
aagttccccc ccgccccgca gctcgcgtcg tgcaggacgt gacaaatgga agtagcacgt    6480
ctcactagtc tcgtgcagat ggacaagcac cgctgagcaa tggaagcggg taggcctttg    6540
gggcagcggc caatagcagc ttggctcctt cgctttctgg gctcagaggc tgggaagggg    6600
tgggtccggg ggcgggctca ggggcgggct caggggcggg gcgggcgcga aggtcctccg    6660
gacccggcat tctgcacgct tcaaaagcgc acgtctgccg cgctgttctc ctcttcctca    6720
tctccgggcc ttcgacctgc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt    6780
ttctgatcga aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat    6840
ctcgtgcttt cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg    6900
ccgatggttt ctacaaagat cgttatgttt atcggcactt gcatcggcc gcgctcccga     6960
ttccggaagt gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc    7020
gtgcacaggg tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc    7080
cggtcgcgga ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg    7140
gcccattcgg accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga    7200
ttgctgatcc ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg    7260
tcgcgcaggc tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc    7320
tcgtgcacgc ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg    7380
```

```
tcattgactg gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct  7440
tctggaggcc gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc  7500
cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac  7560
tctatcagag cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg  7620
acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg  7680
cggccgtctg gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca  7740
gcactcgtcc gagggcaaag gaatagagta tgcagaaatt gatgatctat taaacaataa  7800
agatgtccac taaaatggaa gtttttcctg tcatactttg ttaagaaggg tgagaacaga  7860
gtacctacat tttgaatgga aggattggag ctacgggggt gggggtgggg tgggattaga  7920
taaatgcctg ctctttactg aaggctcttt actattgctt tatgataatg tttcatagtt  7980
ggatatcata atttaaacaa gcaaaaccaa attaagggcc agctcattcc tcccactcat  8040
gatctataga tctatagatc tctcgtggga tcattgtttt tctcttgatt cccactttgt  8100
ggttctaagt actgtggttt ccaaatgtgt cagtttcata gcctgaagaa cgagatcagc  8160
agcctctgtt ccacatacac ttcattctca gtattgtttt gccaagttct aattccatca  8220
gattttagtg gacatcttta ttgtttaata gatcatcaat ttctgcatac taagcttatc  8280
gataccgtcg acctcgaggg ggggcccggt acccagcttt tgttcccttt agtgagggtt  8340
aattgcgcgc ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct  8400
cacaattcca cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaatg  8460
agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct  8520
gtcgtgccag ctgcattaat gaatcggcca cgcgcgggg agaggcggtt tgcgtattgg  8580
gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg tcgttcggc tgcggcgagc  8640
ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg  8700
aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct  8760
ggcgtttttc cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca  8820
gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct  8880
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc  8940
gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt  9000
tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc  9060
cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc  9120
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg  9180
gtggcctaac tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc  9240
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag  9300
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga  9360
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat  9420
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag  9480
ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat  9540
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc  9600
cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat  9660
accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag  9720
ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg  9780
ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc  9840
tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca  9900
acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg  9960
tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc  10020
actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta  10080
ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc  10140
aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg  10200
ttcttcgggg cgaaaactct caaggatctt accgctgttg atccagtt cgatgtaacc  10260
cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc  10320
aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat  10380
actcatactc ttcctttttc aatattattg aagcatttat caggttatt gtctcatgag  10440
cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc  10500
ccgaaaagtg ccac                                                   10514
```

<210> 46
<211> 9626
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
      DNA vector sequence

<400> 46
```
gacggatcgg gagatctccc gatcccctat ggtcgactct cagtacaatc tgctctgatg      60
ccgcatagtt aagccagtat ctgctccctg cttgtgtgtt ggaggtcgct gagtagtgcg     120
cgagcaaaat ttaagctaca acaaggcaag gcttgaccga caattgcatg aagaatctgc     180
ttagggttag gcgttttgcg ctgcttcgcg atgtacgggc cagatatacg cgttgacatt     240
gattattgac tagttattaa tagtaatcaa ttacggggtc attagttcat agcccatata     300
tggagttccg cgttacataa cttacggtaa atggcccgcc tggctgaccg cccaacgacc     360
cccgcccatt gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc     420
attgacgtca atgggtggac tatttacggt aaactgccca cttggcagta catcaagtgt     480
atcatatgcc aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt     540
atgcccagta catgacctta tgggactttc ctacttggca gtacatctac gtattagtca     600
tcgctattac catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg     660
actcacgggg atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc     720
aaaatcaacg ggactttcca aaatgtcgta acaactccgc cccattgacg caaatgggcg     780
gtaggcgtgt acggtgggag gtctatataa gcagagctct ctggctaact agagaaccca     840
ctgcttactg gcttatcgaa attaatacga ctcactatag ggagacccaa gctggctagc     900
gtttaaactt aagcttggta ccgagctcgg atccactagt ccagtgtggt ggaattctgc     960
agatatccag cacagtggct gcggccgcg aaggaagata gccaaggctt actgaggctg    1020
gtggagggag ccactgctgg gctcaccatg gaccgccgga tgtggggggc ccacgtcttc    1080
tgcgtgttga gcccgttacc gaccgtgagt agccagctga gacctctgg cctctgccac    1140
tgggctccag atttgggagc cacagggcca actggagcct ggcggagacc ctactgccct    1200
tctcctgctc tagagtccct ccctacgagc aggtggcagg ctgtggcttg gagagcccct    1260
gacactgggg ctccctccct catctccctc tgcccagcaa tctgtagtga ctctgggtag    1320
tgctattaac atgtccttcc aggggctgag ctgagtcac cttcttgttg cttcccccag    1380
ggcctggctg tgcctctggg ggccacagga actctgtctg ccacaagaca gtgcctcaga    1440
attggtgagg caaagtcagg ctcccctctt agtccctgca gtttgctctt tctctcagct    1500
gtgcacctgt gatgcttcca cacgctcctg tggtaagctt tctggaccct tccttagccc    1560
tatgactccc tgagccactc tccagtttgg tggtttcctt tgaagagggt gtggccaggg    1620
aaggaatgga acctaaggag ttggtggaga ccagctgctt ctgctccagc acttagggga    1680
tccatcatct tctcacgatg ctgagggctg gagaggcttt tggcaaccaa attgtactcc    1740
atgggctaga gcgctggagc acggggcatc aggagagagg gcccttcctg aatgggctgc    1800
agtttcccag taactttcca ggaggttact gcaagtgaca gcagggacca ggctgggagt    1860
cacaaggttc tgggatggtg agggcctgaa gcattcttcc tccgctgcct tattggaggg    1920
tgctgcccct cccgagcacc ttgctgcacc tgctcctgag tgcacaggct ttgagcttga    1980
atcctggacg catgaaacta gccgtgtgaa ctgcaagttc ctcagccttc ctgagtcttg    2040
cttttatctc ttttgttaaa agagaaaata tccgctgcc cgggtcactg tgcattgagt    2100
ggaatctgat ttgtcaaggg ctgggcatag gggtgcaaca caggaaactt cgactgcccc    2160
cctgtctgcc atctgtgtgc actctacagc aaaagcttta tttttagcag gggtacccaa    2220
gatctggtgg tggccattca gcagatggtg ctagtaccaa agctagaact caatcactgt    2280
ttgccagtgt gggcatcatc ccctgcagat accaggaaga gcagcaggca tagatgcaag    2340
gcctgggggt ccagggcgat ggagcctccc ctcaggggtg ggtggggtcc acgatgtgtg    2400
tgggggggtg actgggagcc ccaggatgtg tttcgccagt cctaattcag caggagcatt    2460
tggataaatc ttagaacgat ctatccttcc actcccacca cttccctccc tagaatctct    2520
ccttacatct ggccctatgc ccacacatcc tgagtcatta gtggggcttc tctgaatttc    2580
cctccttctc ttagaggtgt gggcactgtc tgttcatcct tttggcatga ggctgctctc    2640
ctggcatcct ttgccctcag ttgcttcccc aaattggccc tgaaaagttt cccaggcagg    2700
gttggggctt gccagactgt gcagtaggca gaaggatagg cagagaccac agctcttcca    2760
acagcccagg gtcaatatca ggctgatggt gtgagaaagt cccatacagg cccatctgcc    2820
ctcctgctct tccttccccg ctttgtgcgc tcagtctgtt gggagatgta ggggagacgc    2880
aacgggctc tgccctgaga tggacctgag gttactgagt cctcgggagg aagccaagag    2940
ctgagtccgg ggcacccacc aaaataagac ctggagcctg ggtagtgggt gacatctgca    3000
ccctggggcc ttgcagggcc aggcaggaga tggcagagct gcgtccaagc catggaccca    3060
```

```
tggggtcagc agggcgggac aggagtgact tctcctgcca gggtagcact ggtgggggtg   3120
accctcctta ctgggctggc tctccagatt ttctcctatc ccacaaatct gcggcagagc   3180
tctcgggtct ctccacgtgg tcacccacca tctattgtcc ccactctggt ggtcaaagtc   3240
tctactctct gaccccagcc ctggtcttac atctgctaca ctatttgctg tgcctgcatt   3300
tctcccaaag aacactagct gttaattaaa catttgctga catggatcct gctagacatt   3360
gagggtgtgt gtgtgtgggg tgggggtggg ggccatgggg agagaaacat tagtcctgag   3420
ctgatacttc ccaaatacac cttttcccac ttcccttgtc cttgcacatg ctgtgctttc   3480
tgttccagat gctcttctcc ccactgtgtg tgtctggctc taaaaaatgc tcccttgtcc   3540
ccaggtccag ctcaaatgcc acctccccaa tccctgtgag agggacattc cctcccatgg   3600
tacctgtctt gcttctttt gtttttttt tgttttttctg agatggagtc tcactttatc   3660
gcccaggctg gagtacagtg gcacaatctc agctcactgt aacctccgcc tcctgggttc   3720
aagtgattct cctgcctcag cctcctgagt agctgggatt acaggagcgt gtcatcacac   3780
ccagctaatt tttgtatttt tgtagagata gtgtttttgcc acattggtca ggctggtgtt   3840
taactcctga cgtcaggtga tccgcctacc ttggcctccc aaagtgctgg gattacaggc   3900
atgagccact gcgcccggcc cctgtcctgc ttctgatcac agtttgcgca gggtctatta   3960
ggatccagcc cttctccctt ccccagcccc tcctctcctg tgaaagcctg ggggaacaga   4020
catctacact ggttgacttg gatgtgtgtt ttatacaaaa atgtgtatgc tgtataccca   4080
gttcacagtg cagaaagtat gttttgtaaa cttggcttcc cttgtatgta tctgtttggg   4140
ctctggcctg ctgggcgatc agttttggtt ttccttcggt ttggggcggt gggagggaga   4200
agtgctagaa attcctgctg ttcacttcca gtttcctgga gaacttgtta aaatacaaag   4260
tgctggatcc accctagatc tctacatatt cctgcttcca tcccaatgga tctctagctg   4320
gttcttgagg cacactttca gaagtactgc tagaagattc ccaaaaaaac cccaacttgt   4380
ttccagcttc cagagatcat ttctcatggc cactgtgtag cacaagatct gggaggctcc   4440
taaaactata atctttcaag gttgctcagc catcaaggtt atgatccttt catctacaca   4500
tccatgaagc cacacatcca tccttacaca gctctctcac tcatccacca tgcacccatc   4560
catgcagcca gtcatctgtc tatccatcca tccacctccc cagccatata tttgtctgtc   4620
tgtctgtata ctcatccatt ccactagcct tccttccttc ttttctgcct tcctgccttc   4680
ctgccttcct tccctccttc cttttttat gcagccttca cctcattcat gtgctcactc   4740
atccatgtgc tcattcattc atccatcttc tctctcactt actaagctac ccatttgctc   4800
atgtgttcct gcctgtgtgc acccctaaaa gcagagactg agacacaaggt cagaaattta   4860
aaatggtttc cagtggggta aatccaaagt attggacgag ctgcattcct cctggagact   4920
gtagagaata atccatttcc ttgccttttc cttgggttca tgcccttttct ccatcttcag   4980
accagcagtg tggcatcttt aagcggccgc tcgagtctag agggcccgtt taaacccgct   5040
gatcagcctc gactgtgcct tctagttgcc agccatctgt tgtttgcccc tcccccgtgc   5100
cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaaattg   5160
catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg caggacagca   5220
aggggagga ttgggaagac aatagcaggc atgctgggga tgcggtgggc tctatggctt   5280
ctgaggcgga aagaaccagc tggggctcta ggggtatcc ccacgcgccc tgtagcggcg   5340
cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac cgctacactt gccagcgccc   5400
tagcgcccgc tcctttcgct ttcttccctt cctttctcgc cacgttcgcc ggctttcccc   5460
gtcaagctct aaatcggggc atcccttag ggttccgatt tagtgcttta cggcacctcg   5520
accccaaaaa acttgattag ggtgatggt cacgtagtgg gccatcgccc tgatagacgg   5580
tttttcgccc tttgacgttg gagtccacgt tctttaatag tggactcttg ttccaaactg   5640
gaacaacact caaccctatc tcggtctatt cttttgattt ataagggatt ttggggattt   5700
cggcctattg gttaaaaat gagctgattt aacaaaaatt taacgcgaat taattctgtg   5760
gaatgtgtgt cagttagggt gtggaaagtc cccaggctcc ccaggcaggc agaagtatgc   5820
aaagcatgca tctcaattag tcagcaacca ggtgtggaaa gtccccaggc tccccagcag   5880
gcagaagtat gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc   5940
cgcccatccc gcccctaact ccgcccagtt ccgcccattc tccgccccat ggctgactaa   6000
tctttttat ttatgcagag gccgaggccg cctctgcctc tgagctattc cagaagtagt   6060
gaggaggctt ttttggaggc ctaggctttt gcaaaaagct cccgggagct tgtatatcca   6120
ttttcggatc tgatcagcac gtgatgaaaa agcctgaact caccgcgacg tctgtcgaga   6180
agtttctgat cgaaaagttc gacagcgtct ccgacctgat gcagctctcg gagggcgaag   6240
aatctcgtgc tttcagcttc gatgtaggag ggcgtggata tgtcctgcgg gtaaatagct   6300
gcgccgatgg tttctacaaa gatcgttatg tttatcggca ctttgcatcg gccgcgctcc   6360
cgattccgga agtgcttgac attggggaat tcagcgagag cctgacctat tgcatctccc   6420
gccgtgcaca gggtgtcacg ttgcaagacc tgcctgaaac cgaactgccc gctgttctgc   6480
agccggtcgc ggaggccatg gatgcgatcg ctgcggccga tcttagccag acgagcgggt   6540
```

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 47
gagaacccac tgcttactgg ct   22

<210> 48
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 48
gcattgctgt gcagtctgta a   21

<210> 49
<211> 7354
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
      DNA vector sequence

<400> 49
cgcgattgct gatccccatg tgtatcactg gcaaactgtg atggacgaca ccgtcagtgc   60
gtccgtcgcg caggctctcg atgagctgat gctttgggcc gaggactgcc ccgaagtccg   120
gcacctcgtg cacgcggatt tcggctccaa caatgtcctg acggacaatg gccgcataac   180
agcggtcatt gactggagcg aggcgatgtt cggggattcc caatacgagg tcgccaacat   240
cttcttctgg aggccgtggt tggcttgtat ggagcagcag acgcgctact tcgagcggag   300
gcatccggag cttgcaggat cgccgcggct ccgggcgtat atgctccgca ttggtcttga   360
ccaactctat cagagcttgg ttgacggcaa tttcgatgat gcagcttggg cgcagggtcg   420
atgcgacgca atcgtccgat ccggagccgg gactgtcggg cgtacacaaa tcgcccgcag   480
aagcgcggcc gtctggaccg atggctgtgt agaagtactc gccgatagtg gaaaccgacg   540
ccccagcact cgtccgaggg caaaggaata gagtatgcag aaattgatga tctattaaac   600
aataaagatg tccactaaaa tggaagtttt tcctgtcata ctttgttaag aagggtgaga   660
acagagtacc tacattttga atggaaggat tggagctacg ggggtggggg tggggtggga   720
ttagataaat gcctgctctt tactgaaggc tctttactat tgctttatga taatgtttca   780
tagttggata tcataattta aacaagcaaa accaaattaa gggccagctc attcctccca   840
ctcatgatct atagatctat agatctctcg tgggatcatt gtttttctct tgattcccac   900
tttgtggttc taagtactgt ggtttccaaa tgtgtcagtt tcatagcctg aagaacgaga   960
tcagcagcct ctgttccaca tacacttcat tctcagtatt gttttgccaa gttctaattc   1020
catcagattt tagtggacat ctttattgtt taatagatca tcaatttctg catactaagc   1080
ttatcgatac cgtcgacctc gagggggggc ccggtaccca gcttttgttc cctttagtga   1140
gggttaattg cgcgcttggc gtaatcatgg tcatagctgt ttcctgtgtg aaattgttat   1200
ccgctcacaa ttccacacaa catacgagcc ggaagcataa agtgtaaagc ctggggtgcc   1260
taatgagtga gctaactcac attaattgcg ttgcgctcac tgcccgcttt ccagtcggga   1320
aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt   1380
attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg   1440
cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc agggggataac   1500
gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg   1560
ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca   1620
agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc   1680

```
tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc   1740
ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag ttcggtgtag   1800
gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc   1860
ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca   1920
gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg   1980
aagtggtggc ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg   2040
aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct   2100
ggtagcggtg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa   2160
gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa   2220
gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa   2280
tgaagtttta aatcaatcta agtatatat gagtaaactt ggtctgacag ttaccaatgc    2340
ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat agttgcctga   2400
ctccccgtcg tgtagataac tacgatacgg gagggcttac catctggccc cagtgctgca   2460
atgataccgc gagacccacg ctcaccggct ccagatttat cagcaataaa ccagccagcc   2520
ggaagggccg agcgcagaag tggtcctgca actttatccg cctccatcca gtctattaat   2580
tgttgccggg aagctagagt aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc   2640
attgctacag gcatcgtggt gtcacgctcg tcgtttggta tggcttcatt cagctccggt   2700
tcccaacgat caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc   2760
ttcggtcctc cgatcgttgt cagaagtaag ttggccgcag tgttatcact catggttatg   2820
gcagcactgc ataattctct tactgtcatg ccatccgtaa gatgctttc tgtgactggt    2880
gagtactcaa ccaagtcatt ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg   2940
gcgtcaatac gggataatac cgcgccacat agcagaactt aaaagtgct catcattgga    3000
aaacgttctt cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg   3060
taacccactc gtgcacccaa ctgatcttca gcatctttta ctttcaccag cgtttctggg   3120
tgagcaaaaa caggaaggca aaatgccgca aaaaagggaa taaggcgac acggaaatgt     3180
tgaatactca tactcttcct ttttcaatat tattgaagca tttatcaggg ttattgtctc   3240
atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt tccgcgcaca   3300
tttccccgaa aagtgccacc taaattgtaa gcgttaatat tttgttaaaa ttcgcgttaa   3360
atttttgtta atcagctca tttttttaacc aataggccga atcggcaaa atcccttata    3420
aatcaaaaga atagaccgag atagggttcc gtgttgttcc agtttggaac aagagtccac   3480
tattaaagaa cgtggactcc aacgtcaaag ggcgaaaaac cgtctatcag ggcgatggcc   3540
cactacgtga accatcaccc taatcaagtt ttttggggtc gaggtgccgt aaagcactaa   3600
atcggaaccc taaagggagc ccccgattta gagcttgacg gggaaagccg cgaacgtgg    3660
cgagaaagga agggaagaaa gcgaaaggag cgggcgctag ggcgctggca agtgtagcgg   3720
tcacgctgcg cgtaaccacc acacccgccg cgcttaatgc gccgctacag ggcgcgtccc   3780
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat   3840
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt   3900
tttcccagtc acgacgttgt aaaacgacgg ccagtgagcg cgcgtaatac gactcactat   3960
agggcgaatt ggagctccac cgcggtgcgg ccgggccatg caggccagac atgataagat   4020
acattgatga gtttggacaa accacaacta gaatgcagtg aaaaaaatgc tttatttgtg   4080
aaatttgtga tgctattgct ttatttgtaa ccattataag ctgcaataaa caagttgtcg   4140
agttacttgt acagctcgtc catgccgaga gtgatcccgg cggcggtcac gaactccagc   4200
aggaccatgt gatcgcgctt ctcgttgggg tctttgctca gggcggactg gtagctcagg   4260
tagtggttgt cgggcagcag cacggggccg tcgccgatgg gggtgttctg ctggtagtgg    4320
tcggcgagct gcacgctgcc gtcctcgatg ttgtggcgga tcttgaagtt caccttgatg   4380
ccgttcttct gcttgtcggc catgatatag acgttgtggc tgttgtagtt gtactccagc   4440
ttgtgcccca ggatgttgcc gtcctccttg aagtcgatgc ccttcagctc gatgcggttc   4500
accagggtgt cgccctcgaa cttcacctcg gcgcgggtct tgtagttgcc gtcgtccttg   4560
aagaagatgg tgcgctcctg gacgtagcct tcgggcatgg cggacttgaa gaagtcgtgc   4620
tgcttcatgt ggtcggggta gcggcgaag cactgcaggc cgtagccgaa ggtggtcacg    4680
agggtgggcc agggcacggg cagcttgccg gtggtgcaga tgaacttcag ggtcagcttg   4740
ccgtaggtgg catcgccctc gccctcgccg gacacgctga cttgtggcc gtttacgtcg     4800
ccgtccagct cgaccaggat gggcaccacc ccggtgaaca gctcctcgcc cttgctcacc   4860
atggtggcgt tagcttgatt tgacagtggc tgggggttgc gccgccgggt tttataggaa    4920
gccacagcgg ccactcgagc cataaaaggc aactttagga acggcgggg gtgattggat     4980
tcgagtcgtt tattcaccgg ccttgccgca cagtgcagca ttttttttacc ccctctcccc   5040
tcctttgcg ggggaaaaaa aaaaaaaaaa aaaaaaaagg agaagagaaa aaaagcgagc     5100
gagagagaaa gcgagattga ggaagaggat gaagagtttg gcgatgggtg ctggttccgt   5160
```

```
aggcccagat ggacaagaat agccccccgcc cttgcggaca gtatcccatt cagtgactca   5220
gatcagtcaa gcggccgctc tagacggcga tacagtcaac tgtctttgac cttttgttact   5280
actctcttcc gatgatgatg tcgcacttat tctatgctgt ctcaatgtta gaggcatatc   5340
agtctccact gaagccaatc tatctgtgac ggcatcttta ttcacattat cttgtacaaa   5400
taatcctgtt aacaatgctt ttatatcctg taaagaatcc attttcaaaa tcatgtcaag   5460
gtcttctcga ggaaaaatca gtagaaatag ctgttccagt ctttctagcc ttgattccac   5520
ttctgtcaga tgtgccctag tcagcggaga cctttttggtt ttgggagagt agcgacactc   5580
ccagttgttc ttcagacact tggcgcactt cggttttttct ttggagcact tgagcttttt   5640
aagtcggcaa atatcgcatg cttgttcgat agaagacagt agcttcatgg tggcggatcc   5700
cgataagcca gtaagcagtg ggttctctag ttagccagag agctctgctt atatagacct   5760
cccaccgtac acgcctaccg cccatttgcg tcaatggggc ggagttgtta cgacattttg   5820
gaaagtcccg ttgatttttgg tgccaaaaca aactcccatt gacgtcaatg gggtggagac   5880
ttggaaatcc ccgtgagtca aaccgctatc cacgcccatt gatgtactgc caaaaccgca   5940
tcaccatggt aatagcgatg actaatacgt agatgtactg ccaagtagga aagtcccata   6000
aggtcatgta ctgggcataa tgccaggcgg gccatttacc gtcattgacg tcaatagggg   6060
gcgtacttgg catatgatac acttgatgta ctgccaagtg ggcagtttac cgtaaatact   6120
ccacccattg acgtcaatgg aaagtcccta ttggcgttac tatgggaaca tacgtcatta   6180
ttgacgtcaa tgggcggggg tcgttgggcg gtcagccagg cgggccattt accgtaagtt   6240
atgtaacgcg gaactccata tatgggctat gaactaatga ccccgtaatt gattactatt   6300
aataactagt caataatcaa tgtcaacgcg tatatctggc ccgtacatcg cgaagcagcg   6360
caaaacgcct aaccctaccc gggctgcagg aattctaccg ggtaggggag gcgcttttcc   6420
caaggcagtc tggagcatgc gctttagcag ccccgctggc acttggcgca tcacaagtgg   6480
cctctggcct cgcacacatt ccacatccac cggtagcgcc aaccggctcc cttctttggt   6540
ggccccttcg cgccaccttc tactcctccc ctagtcagga agttcccccc cgccccgcag   6600
ctcgcgtcgt gcaggacgtg acaaatggaa gtagcacgtc tcactagtct cgtgcagatg   6660
gacaagcacc gctgagcaat ggaagcgggt aggcctttgg ggcagcggcc aatagcagct   6720
tggctccttc gctttctggg ctcagaggct gggaaggggt gggtccgggg gcgggctcag   6780
gggcgggctc aggggcgggg cgggcgcgaa ggtcctccgg acccggcatt ctgcacgctt   6840
caaaagcgca cgtctgccgc gctgttctcc tcttcctcat ctccgggcct tcgacctgca   6900
tgaaaaagcc tgaactcacc gcgacgtctg tcgagaagtt tctgatcgaa aagttcgaca   6960
gcgtctccga cctgatgcag ctctcggagg gcgaagaatc tcgtgctttc agcttcgatg   7020
taggagggcg tggatatgtc ctgcgggtaa atagctgcgc cgatggtttc tacaaagatc   7080
gttatgttta tcggcacttt gcatcggccg cgctcccgat tccggaagtg cttgacattg   7140
gggaattcag cgagagcctg acctattgca tctcccgccg tgcacagggt gtcacgttgc   7200
aagacctgcc tgaaaccgaa ctgcccgctg ttctgcagcc ggtcgcggag gccatggatg   7260
cgatcgctgc ggccgatctt agccagacga gcgggttcgg cccattcgga ccgcaaggaa   7320
tcggtcaata cactacatgg cgtgatttca tatg                                7354
```

<210> 50
<211> 10514
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
     DNA vector sequence

<400> 50

```
ctaaattgta agcgttaata ttttgttaaa attcgcgtta aattttttgtt aaatcagctc    60
atttttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga   120
gatagggttg agtgttgttc cagtttggaa caagagtcca ctattaaaga acgtggactc   180
caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc   240
ctaatcaagt tttttggggt cgaggtgccg taaagcacta aatcggaacc ctaaagggag   300
cccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa   360
agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac   420
cacacccgcc gcgcttaatg cgccgctaca gggcgcgtcc cattcgccat tcaggctgcg   480
caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg   540
```

```
gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt cacgacgttg    600
taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaat tggagctcca    660
ccgcggtgcg gccgggccat gcaggccaga catgataaga tacattgatg agtttggaca    720
aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc    780
tttatttgta accattataa gctgcaataa acaagttgtc gagttacttg tacagctcgt    840
ccatgccgag agtgatcccg gcggcggtca cgaactccag caggaccatg tgatcgcgct    900
tctcgttggg gtctttgctc agggcggact ggtagctcag gtagtggttg tcgggcagca    960
gcacggggcc gtcgccgatg ggggtgttct gctggtagtg gtcggcgagc tgcacgctgc   1020
cgtcctcgat gttgtggcgg atcttgaagt tcaccttgat gccgttcttc tgcttgtcgg   1080
ccatgatata gacgttgtgg ctgttgtagt tgtactccag cttgtgcccc aggatgttgc   1140
cgtcctcctt gaagtcgatg cccttcagct cgatgcggtt caccagggtg tcgccctcga   1200
acttcacctc ggcgcgggtc ttgtagttgc cgtcgtcctt gaagaagatg gtgcgctcct   1260
ggacgtagcc ttcgggcatg gcggacttga agaagtcgtg ctgcttcatg tggtcggggt   1320
agcgggcgaa gcactgcagg ccgtagccga aggtggtcac gagggtgggc cagggcacgg   1380
gcagcttgcc ggtggtgcag atgaacttca gggtcagctt gccgtaggtg gcatcgccct   1440
cgccctcgcc ggacacgctg aacttgtggc cgtttacgtc gccgtccagc tcgaccagga   1500
tgggcaccac cccggtgaac agctcctcgc ccttgctcac catggtggcg ttagcttgat   1560
ttgacagtgg ctgggggttg cgccgccggg ttttatagga agccacagcg gccactcgag   1620
ccataaaagg caactttagg aacggcgggg ggtgattgga ttcgagtcgt ttattcaccg   1680
gccttgccgc acagtgcagc attttttac cccctctccc ctccttttgc gggggaaaaa   1740
aaaaaaaaaa aaaaaaaag gagaagagaa aaaaagcgag cgagagagaa agcgagattg   1800
aggaagagga tgaagagttt ggcgatgggt gctggttccg taggcccaga tggacaagaa   1860
tagcccccgc ccttgcggac agtatcccat tcagtgactc agatcagtca agcggccgcg   1920
tctaacctag gagcccacca cagacactat actataacca cactgaattt tctctcctat   1980
ctcaaataat cctattcttg ttcatgcctc agtgccttat aacaagcaga tttctcggcc   2040
agacatggtg gcttgcgcct gtaatcccag cactttggcc aagtcgggca gctcacttga   2100
ggtcaggagt ttgaattcag cctgtccaat atggtgaaac cccatctcta ctaaaaatac   2160
aaaaattagc tgggcatggt ggtgtgtgcc tgtagtccca gctactcagg agactgaggc   2220
aggagaattg cttgaactcg ggaagcagag gttgcagtga gccaagatca tgccattgca   2280
ctccagcctg ggcaacagag cgagactccg tctcaaacaa aacaaaacaa aaacaaaaaa   2340
acaagcagat ctctcaactt ggaatgctgc tgcttctcca tgccaactgg aaaacttctt   2400
tttcttttag actcaactca aatatgatcc tccctgaagt ctctccatct ttcctagaca   2460
caattagcaa ccagttcctc tgtgttcaaa cagcacttcg tgtgtatgtg tatgtatgtg   2520
tatttatcat tattattatc acctttgatc actctccctg aaggtaatag ccaggcactt   2580
agtatagagt aggtatcaac aattgcttat tgaaggaata aatgggtgaa tgacaacttt   2640
gcaaagtttc aattttaat ctattgattt gaaagaatta ttatacatta ttttaaatat   2700
cctttctagt tctaatatct tatgtctata gtagttcgac agttactta aatcagatat   2760
taaatgattt tcattttaca aagaacagac ctgcctttct caatgtcaca agactattgt   2820
agagaatagc agtgaaataa gtacatgtga aaatacttgg caaagtaata gtgaaaatag   2880
agctgatctc acaggaatgt tttgaaaatt aaaacaccgt atgtaaaata ggtaatctaa   2940
tctgtgggca catatttaaa ctctcaataa atattaactt tgattattta aaaaatacct   3000
ggtagctttt ttgtgaattc tacaagaacc tgtacatgat tggttgccat ttccgtcaaa   3060
atgagaaaat tttcttctgc actgaattct tctttttaact aaattttaag aaaaaaacca   3120
atgattaaaa agtagataag aaaatgataa tataaaacca ttagactaaa gaaaaactag   3180
gtttagatga taaaatttga aaacccatct tactatttct cagatctttg gagggcaatt   3240
accattatca agataggtgg gagacttcat actcttgaat tgtttaatgt tgctgagagt   3300
ttaaaagtct aacttgaatt ctcagctcta atcatttcaa agtaacatat tcacatataa   3360
attatacctc caaggcaata aaaactccag aaaatattaa ctttaaaccc acatattttc   3420
agatattata catacaattt tatttgttat ttcctgaggc atcctctgtt tgttatatga   3480
atccataata aaatgtagaa gagtctgttg atctggggtg agttcagttt tctcctacac   3540
tggcaagaaa ataaaaggtg ttagggaaga tgtgtataaa gccttctttc tatatgtgca   3600
ttcatctgga gggacctaat agcagaaatc ataaaaaaaa aaaggcataa atacagcagg   3660
aactatgagt tctaagagag tgtataattt ctttctctgg gtcctcaggc aggccatatc   3720
acatttcact gaattttcct atctataaaa tatgaataac aaaatattcc acttgtgtat   3780
ttcttaggaa taaatggtag tctggatatg aagactattg agcccattca tattgaaggt   3840
cactgaaagt tcaagcttta attataacta atgaaaggtt ttcatatagt atgactacta   3900
tgaatactat attcatatac tattctacat gaattccacc tggtgaagtc atctgagagt   3960
attcatgtgg aggagggatc aacaaactgt ggcccatgga ccaaatccat tggattttat   4020
```

```
aaagtttat  tggaacacag  tcacactcat  ttgtttatgt  attgtctctg  ctactttgc   4080
actaaaatgg  cagagttgag  tagctgttgt  agataccatg  tggccaaagc  cctttaaaga  4140
aaaaaagttt  gtcaacccct  aaaatagaga  gtcaacaatc  tggacaattc  agaccccaaa  4200
catctctttc  tccctctaca  gcatcttttg  aagagtgtga  ataattatgt  caactctttg  4260
ttttcttata  ccaaagtgtt  actgatgaag  cactatggca  aaataatgct  ttcatataaa  4320
atactttttc  ttaaataaat  acaaaaatac  gtatttgtaa  aactaaatgt  ggataaagtt  4380
ttctttgtca  ataatgtaat  aaatattgag  aagggtgtgg  tttggacaat  catagcttac  4440
tggttttttca gctaatggca  aaataaaatt  ccatcaaatg  tcctgtgcat  tatcacaatg  4500
tgtgtgtcta  aacaccttgg  atgctgttcc  cctattttgt  cctaagttgc  tcaatggatt  4560
tcttgagaaa  tttgtaaccc  cagtgggtag  gacttttctc  tattcatcca  catacgtgta  4620
ttcaccctgt  tttactcatt  tataaggatt  aatccacagc  ctctattgag  accatgttct  4680
tcccagtttg  agactgtctt  ctcaaaaaga  gggaaaaaat  ggtcatcaat  actcagttgg  4740
aaaaaatcca  ttttatacac  caaatcattt  ctgctggcaa  acactaggta  gtaggactgg  4800
aatggagtca  gagagaagga  aaactttcac  tttttagttt  atgcattgct  actatttgaa  4860
tgtttacatc  aaagcatgta  ttaggaatcc  attttaatga  acaattggac  tagaaattca  4920
gctgtgcctg  tggctttaat  ttgaaagcct  cctggccagc  tcactatgta  cctagaaact  4980
cagttcctgc  cagtcttggc  agacaccatt  gcatattatt  accctgcatg  actttgttgt  5040
cgaggtcact  tgtcgcaagt  cacgaccttc  actgtcttct  ctagacggcg  atacagtcaa  5100
ctgtctttga  cctttgttac  tactctcttc  cgatgatgat  gtcgcactta  ttctatgctg  5160
tctcaatgtt  agaggcatat  cagtctccac  tgaagccaat  ctatctgtga  cggcatcttt  5220
attcacatta  tcttgtacaa  ataatcctgt  taacaatgct  tttatatcct  gtaaagaatc  5280
cattttcaaa  atcatgtcaa  ggtcttctcg  aggaaaaatc  agtagaaata  gctgttccag  5340
tctttctagc  cttgattcca  cttctgtcag  atgtgcccta  gtcagcggag  accttttggt  5400
tttgggagag  tagcgacact  cccagttgtt  cttcagacac  ttggcgcact  tcggttttc   5460
tttggagcac  ttgagctttt  taagtcggca  aatatcgcat  gcttgttcga  tagaagacag  5520
tagcttcatg  gtggcggatc  ccgataagcc  agtaagcagt  gggttctcta  gttagccaga  5580
gagctctgct  tatatagacc  tcccaccgta  cacgcctacc  gcccatttgc  gtcaatgggg  5640
cggagttgtt  acgacatttt  ggaaagtccc  gttgatttttg  gtgccaaaac  aaactcccat  5700
tgacgtcaat  ggggtggaga  cttggaaatc  cccgtgagtc  aaaccgctat  ccacgcccat  5760
tgatgtactg  ccaaaaccgc  atcaccatgg  taatagcgat  gactaatacg  tagatgtact  5820
gccaagtagg  aaagtcccat  aaggtcatgt  actgggcata  atgccaggcg  ggccatttac  5880
cgtcattgac  gtcaataggg  ggcgtacttg  gcatatgata  cacttgatgt  actgccaagt  5940
gggcagttta  ccgtaaatac  tccacccatt  gacgtcaatg  gaaagtccct  attggcgtta  6000
ctatgggaac  atacgtcatt  attgacgtca  atgggcgggg  gtcgttgggc  ggtcagccag  6060
gcgggccatt  taccgtaagt  tatgtaacgc  ggaactccat  atatgggcta  tgaactaatg  6120
accccgtaat  tgattactat  taataactag  tcaataatca  atgtcaacgc  gtatatctgg  6180
cccgtacatc  gcgaagcagc  gcaaaacgcc  taaccctacc  cgggctgcag  gaattctacc  6240
gggtagggga  ggcgcttttc  ccaaggcagt  ctggagcatg  cgctttagca  gccccgctgg  6300
cacttggcgc  atcacaagtg  gcctctggcc  tcgcacacat  tccacatcca  ccggtagcgc  6360
caaccggctc  ccttctttgg  tggccccttc  gcgccacctt  ctactcctcc  cctagtcagg  6420
aagttccccc  ccgccccgca  gctcgcgtcg  tgcaggacgt  gacaaatgga  agtagcacgt  6480
ctcactagtc  tcgtgcagat  ggacaagcac  cgctgagcaa  tggaagcggg  taggcctttg  6540
gggcagcggc  caatagcagc  ttggctcctt  cgctttctgg  gctcagaggc  tgggaagggg  6600
tgggtccggg  ggcgggctca  ggggcgggct  caggggcggg  gcgggcgcga  aggtcctccg  6660
gacccggcat  tctgcacgct  tcaaaagcgc  acgtctgccg  cgctgttctc  ctcttcctca  6720
tctccgggcc  ttcgacctgc  atgaaaaagc  ctgaactcac  cgcgacgtct  gtcgagaagt  6780
ttctgatcga  aaagttcgac  agcgtctccg  acctgatgca  gctctcggag  ggcgaagaat  6840
ctcgtgcttt  cagcttcgat  gtaggagggc  gtggatatgt  cctgcgggta  aatagctgcg  6900
ccgatggttt  ctacaaagat  cgttatgttt  atcggcactt  tgcatcggcc  gcgctcccga  6960
ttccggaagt  gcttgacatt  ggggaattca  gcgagagcct  gacctattgc  atctcccgcc  7020
gtgcacaggg  tgtcacgttg  caagacctgc  ctgaaaccga  actgcccgct  gttctgcagc  7080
cggtcgcgga  ggccatggat  gcgatcgctg  cggccgatct  tagccagacg  agcgggttcg  7140
gcccattcgg  accgcaagga  atcggtcaat  acactacatg  gcgtgatttc  atatgcgcga  7200
ttgctgatcc  ccatgtgtat  cactggcaaa  ctgtgatgga  cgacaccgtc  agtgcgtccg  7260
tcgcgcaggc  tctcgatgag  ctgatgcttt  gggccgagga  ctgccccgaa  gtccggcacc  7320
tcgtgcacgc  ggatttcggc  tccaacaatg  tcctgacgga  caatggccgc  ataacagcgg  7380
tcattgactg  gagcgaggcg  atgttcgggg  attcccaata  cgaggtcgcc  aacatcttct  7440
tctggaggcc  gtggttggct  tgtatggagc  agcagacgcg  ctacttcgag  cggaggcatc  7500
```

```
cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac    7560
tctatcagag cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg    7620
acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg    7680
cggccgtctg gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca    7740
gcactcgtcc gagggcaaag gaatagagta tgcagaaatt gatgatctat taaacaataa    7800
agatgtccac taaaatggaa gtttttcctg tcatactttg ttaagaaggg tgagaacaga    7860
gtacctacat tttgaatgga aggattggag ctacgggggt ggggtgggg tgggattaga    7920
taaatgcctg ctcttactg aaggctcttt actattgctt tatgataatg tttcatagtt    7980
ggatatcata atttaaacaa gcaaaaccaa attaagggcc agctcattcc tcccactcat    8040
gatctataga tctatagatc tctcgtggga tcattgtttt tctcttgatt cccactttgt    8100
ggttctaagt actgtggttt ccaaatgtgt cagtttcata gcctgaagaa cgagatcagc    8160
agcctctgtt ccacatacac ttcattctca gtattgtttt gccaagttct aattccatca    8220
gattttagtg gacatcttta ttgtttaata gatcatcaat ttctgcatac taagcttatc    8280
gataccgtcg acctcgaggg ggggcccggt acccagcttt tgttcccttt agtgagggtt    8340
aattgcgcgc ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct    8400
cacaattcca cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaatg    8460
agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct    8520
gtcgtgccag ctgcattaat gaatcggcca cgcgcgggg agaggcggtt tgcgtattgg    8580
gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg tcgttcggc tgcggcgagc    8640
ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg    8700
aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct    8760
ggcgtttttc cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca    8820
gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct    8880
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc    8940
gggaagcgtg cgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt    9000
tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc    9060
cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc    9120
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg    9180
gtggcctaac tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc    9240
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag    9300
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga    9360
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat    9420
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag    9480
ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat    9540
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc    9600
cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat    9660
accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag    9720
ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg    9780
ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc    9840
tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca    9900
acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg    9960
tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc    10020
actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta    10080
ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc    10140
aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg    10200
ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc    10260
cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc    10320
aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg cgacacgga aatgttgaat    10380
actcatactc ttcctttttc aatattattg aagcatttat cagggttatt gtctcatgag    10440
cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc    10500
ccgaaaagtg ccac                                                      10514
```

<210> 51
<211> 10247
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
      DNA molecule

<220>
<221> modified_base
<222> (2626)..(2626)
<223> unsequenced nucleotide

<400> 51
```
ctagaactag tggatccgga taagccagta agcagtgggt tctctagtta gccagagagc    60
tctgcttata tagacctccc accgtacacg cctaccgccc atttgcgtca atggggcgga   120
gttgttacga cattttggaa agtcccgttg attttggtgc caaaacaaac tcccattgac   180
gtcaatgggg tggagacttg gaaatccccg tgagtcaaac cgctatccac gcccattgat   240
gtactgccaa aaccgcatca ccatggtaat agcgatgact aatacgtaga tgtactgcca   300
agtaggaaag tcccataagg tcatgtactg ggcataatgc caggcgggcc atttaccgtc   360
attgacgtca ataggggcg tacttggcat atgatacact tgatgtactg ccaagtgggc   420
agtttaccgt aaatactcca cccattgacg tcaatggaaa gtccctattg gcgttactat   480
gggaacatac gtcattattg acgtcaatgg gcgggggtcg ttgggcggtc agccaggcgg   540
gccatttacc gtaagttatg taacgcggaa ctccatatat gggctatgaa ctaatgaccc   600
cgtaattgat tactattaat aactagtcaa taatcaatgt caaccccggg ctgcaggaat   660
tctaccgggt aggggaggcg ctttttccaa ggcagtctgg agcatgcgct ttagcagccc   720
cgctggcact tggcgcatca caagtggcct ctggcctcgc acacattcca catccaccgg   780
tagcgccaac cggctccctt ctttggtggc cccttcgcgc caccttctac tcctcccta   840
gtcaggaagt tcccccccgc cccgcagctc gcgtcgtgca ggacgtgaca aatggaagta   900
gcacgtctca ctagtctcgt gcagatggac aagcaccgct gagcaatgga agcgggtagg   960
cctttggggc agcggccaat agcagcttgg ctccttcgct ttctgggctc agaggctggg  1020
aaggggtggg tccggggggcg ggctcagggg cgggctcagg ggcggggcgg gcgcgaaggt  1080
cctccggacc cggcattctg cacgcttcaa aagcgcacgt ctgccgcgct gttctcctct  1140
tcctcatctc cgggccttcg acctgcatga aaaagcctga actcaccgcg acgtctgtcg  1200
agaagtttct gatcgaaaag ttcgacagcg tctccgacct gatgcagctc tcggagggcg  1260
aagaatctcg tgctttcagc ttcgatgtag gagggcgtgg atatgtcctg cgggtaaata  1320
gctgcgccga tggtttctac aaagatcgtt atgtttatcg gcactttgca tcggccgcgc  1380
tcccgattcc ggaagtgctt gacattgggg aattcagcga gagcctgacc tattgcatct  1440
cccgccgtgc acagggtgtc acgttgcaag acctgcctga aaccgaactg cccgctgttc  1500
tgcagccggt cgcggaggcc atggatgcga tcgctgcggc cgatcttagc cagacgagcg  1560
ggttcggccc attcggaccg caaggaatcg gtcaatacac tacatggcgt gatttcatat  1620
gcgcgattgc tgatccccat gtgtatcact ggcaaactgt gatggacgac accgtcagtg  1680
cgtccgtcgc gcaggctctc gatgagctga tgctttgggc cgaggactgc cccgaagtcc  1740
ggcacctcgt gcacgcggat ttcggctcca acaatgtcct gacggacaat ggccgcataa  1800
cagcggtcat tgactggagc gaggcgatgt tcggggattc ccaatacgag gtcgccaaca  1860
tcttcttctg gaggccgtgg ttggcttgta tggagcagca gacgcgctac ttcgagcgga  1920
ggcatccgga gcttgcagga tcgccgcggc tccgggcgta tatgctccgc attggtcttg  1980
accaactcta tcagagcttg gttgacggca atttcgatga tgcagcttgg gcgcagggtc  2040
gatgcgacgc aatcgtccga tccggagccg ggactgtcgg gcgtacacaa atcgcccgca  2100
gaagcgcggc cgtctggacc gatggctgtg tagaagtact cgccgatagt ggaaaccgac  2160
gccccagcac tcgtccgagg gcaaaggaat gcctgagaaa ggaagtgagc tgtaaaggct  2220
gagctctctc tctgacgtat gtagcctctg gttagcttcg tcactcactg ttcttgactc  2280
agcatggcaa tctgatgaaa tcccagctgt aagtctgcat aaattgatga tctattaaac  2340
aataaagatg tccactaaaa tggaagtttt ttactgtcat actttgttaa gaagggtgag  2400
aacagagtac ctacattttg aatggaagga ttggagctac ggggtgggg gtgggggtgg  2460
gattagataa atgcctgctc tttactgaag gctctttact attgctttat gataatgttt  2520
catagttgga tatcataatt taaacaagca aaaccaaatt aagggccagc tcattcctcc  2580
actcatgatc tatagatcta tagatctctc gtgggatcat tgtttntctg atccactgga  2640
agcttatcga taccgtcgac ctcgagggg ggcccggtac ccagcttttg ttccctttag  2700
tgagggttaa ttgcgcgctt ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt  2760
tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa agcctggggt  2820
gcctaatgag tgagctaact cacattaatt gcgttgcgct cactgcccgc tttccagtcg  2880
```

```
ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag aggcggtttg   2940
cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg   3000
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga atcaggggat   3060
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc   3120
gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa aaatcgacgc   3180
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga   3240
agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt   3300
ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg   3360
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc   3420
gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg   3480
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc   3540
ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg   3600
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc   3660
gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct   3720
caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt   3780
taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa   3840
aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa   3900
tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc   3960
tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct   4020
gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca   4080
gccggaaggg ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt   4140
aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttgtt   4200
gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc   4260
ggttcccaac gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa agcggttagc   4320
tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatc actcatggtt   4380
atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt ttctgtgact   4440
ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc   4500
ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt   4560
ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag atccagttcg   4620
atgtaaccca ctcgtgcacc caactgatct tcagcatctt ttactttcac cagcgtttct   4680
gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacgaaa   4740
tgttgaatac tcatactctt cctttttcaa tattattgaa gcatttatca gggttattgt   4800
ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc   4860
acatttcccc gaaaagtgcc acctaaattg taagcgttaa tattttgtta aaattcgcgt   4920
taaattttg ttaaatcagc tcattttta accataggc cgaaatcggc aaaatccctt   4980
ataaatcaaa agaatagacc gagatagggt tgagtgttgt tccagtttgg aacaagagtc   5040
cactattaaa gaacgtggac tccaacgtca aagggcgaaa aaccgtctat cagggcgatg   5100
gcccactacg tgaaccatca ccctaatcaa gttttttggg gtcgaggtgc cgtaaagcac   5160
taaatcggaa ccctaaaggg agccccgat ttagagcttg acggggaaag ccggcgaacg   5220
tggcgagaaa ggaagggaag aaagcgaaag gagcgggcgc tagggcgctg gcaagtgtag   5280
cggtcacgct gcgcgtaacc accacacccg ccgcgcttaa tgcgccgcta cagggcgcgt   5340
cccattcgcc attcaggctg cgcaactgtt gggaagggcg atcggtgcgg gcctcttcgc   5400
tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg attaagttgg gtaacgccag   5460
ggttttccca gtcacgacgt tgtaaaacga cggccagtga gcgcgcgtaa tacgactcac   5520
tataggcga attggagctc caccgcggtg cggccgggcc atgcaggcca cgacatgata   5580
agatacattg atgagtttgg acaaaccaca actagaatgc agtgaaaaaa atgctttatt   5640
tgtgaaattt gtgatgctat tgctttattt gtaaccatta taagctgcaa taaacaagtt   5700
cttgtacagc tcgtccatgc cgagagtgat cccggcggcg gtcacgaact ccagcaggac   5760
catgtgatcg cgcttctcgt tggggtcttt gctcagggcg gactggtagc tcaggtagtg   5820
gttgtcgggc agcagcacgg ggccgtcgcc gatggggggtg ttctgctggt agtggtcggc   5880
gagctgcacg ctgccgtcct cgatgttgtg gcggatcttg aagttcacct tgatgccgtt   5940
cttctgcttg tcggccatga tatagacgtt gtggctgttg tagttgtact ccagcttgtg   6000
ccccaggatg ttgccgtcct ccttgaagtc gatgcccttc agctcgatgc ggttcaccag   6060
ggtgtcgccc tcgaacttca cctcggcgcg ggtcttgtag ttgccgtcgt ccttgaagaa   6120
gatggtgcgc tcctggacgt agccttcggg catggcggac ttgaagaagt cgtgctgctt   6180
catgtggtcg gggtagcggg cgaagcactg caggccgtag ccgaaggtgg tcacgagggt   6240
gggccagggc acgggcagct tgccggtggt gcagatgaac ttcagggtca gcttgccgta   6300
ggtggcatcg ccctcgccct cgccggacac gctgaacttg tggccgttta cgtcgccgtc   6360
```

```
cagctcgacc aggatgggca ccaccccggt gaacagctcc tcgcccttgc tcaccatggt    6420
ggcgttagct tgatttgaca gtggctgggg gttgcgccgc cgggtttat  aggaagccac    6480
agcggccact cgagccataa aaggcaactt taggaacggc gggggtgat  tggattcgag    6540
tcgtttattc accggccttg ccgcacagtg cagcatttt  ttacccctc  tcccctcctt    6600
ttgcggggga aaaaaaaaa  aaaaaaaaaa aaaggagaag agaaaaaaag cgagcgagag    6660
agaaagcgag attgaggaag aggatgaaga gtttggcgat gggtgctggt tccgtaggcc    6720
cagatggaca agaatagccc ccgcccttgc ggacagtatc ccattcagtg actcagatca    6780
gatcaagcgg ccgccagtgt gatggatatc tgcagaattc cagcacactg gcggccgtta    6840
ctagtggatc ctcaatcagt ccatcacctg ggtactttt  ctacccaatg ctgaaagtga    6900
atatttgaga aactatgttt ctaactggaa gaattctgaa actatgttct tctacaacat    6960
gtagacgatg tgtcacaatg aaaatgaaag ttccttctga agagctgtaa ccctacatag    7020
cctaggatgc catttatact ttctctagcc cagaggctct gtatttctgc tctcaggcta    7080
cattactaca taaccatcaa accaactgcc cgaaattgct gagcaattga tcataacata    7140
cagacctgat tcagatatat aaattacatt tttatttta  tcgtaccaaa acatacaggt    7200
gaaaactttc aaattcaggc aactgacatg ccaggtgcct acctatctgt gaagaccaga    7260
gcactgtaag acacggaata gctggttgaa aagtgaaaaa tccagtgagt taggttgttt    7320
gattatacag aatcatagaa ttttcaaagt gaaaggacat ggtgatcact gatttcagtg    7380
actatccaac aacattagca agagcccta  gggttcccta ccagttcctt cagaagcctc    7440
cttcaaagct gtgagtgcat ggccaaaggg gaagctacat tcatcaacca gactaatcag    7500
agaccaggga gcagaaataa ggttaacacg gaatggagta catttgcttt gggggacgtt    7560
catgtttct  tgcgcatact gaaaacagct gggtgctgat tttggtataa cactcagagc    7620
ctcttttgat aaaaaggcat tctatatgaa tgaaaaccct gttggcttca cattggtcaa    7680
agcaatggat tttaaatagt gactcaactt cctagaaaac tttgtattta gaaatgtttt    7740
cttcatacag ctaacaggca aaactaatca aaaccttgtt gttagtctga tgtttatttt    7800
ctgtgcagca aacaaactcc cataaaaacc tctgaggttt tcttagttat gtttgcaatt    7860
ttttaagtgt tgaactgcct aagtctctga tgtcatgaca cttaattcat atatttctt    7920
gcttttgcta gaactgaaca aacactccct atcaagattg accctctcta gaaataaaat    7980
aaattgtcat atgaagtcaa aatatgacct atcaagttta tgatagcttt ctttgacatt    8040
tatgggagga gaaaattata ccatccatac tacataggt  ggggtttaaa taatatttgc    8100
ctataaacaa taagcaattg cagattgatt ttaagtaaac gtatgcaatg cctgtaagca    8160
acaggaaata tatttacat  caatgtatgc actgcctgta agtaacagga attatatttt    8220
acatgaatgt gtgaagtgat aatttttaa  taggcaaaaa ttagtgttaa aagcggaact    8280
ggcatgcttt cctaggctgt ctcagtgata atgacctcaa agtcaatgat ttcttcatga    8340
agatgaaata agtttattgc aacttgattg catcacctaa aatacaaccc ttgacttgga    8400
agcttgccct tatgttttat gaggtctgaa ataaaaagga aactgatgga actgaaccag    8460
caaattattt tgcaatccat gactacagtt gtatacagca gatgtagtca attgttgagg    8520
aaagttggac caaatttcca tcctgttcag tactgctgat agctctattc aaaatacatt    8580
cttgccttta gaggaactga cccattagta atcatacaga tagaattttg tgaactattt    8640
ttaaagggta ggttattaat caatgagtga ctctcaaata tcttgatggt ttccactgtg    8700
aaaaccaagt atctgcatga tagagcaaaa atttaatttt ctcagatgct ttagagtaca    8760
tatagatgaa gactgccttc ttactagatc agttttgaga gagatgaaaa caaaagaaag    8820
aaaaaatagg atagctgatg aaatgaaact tcaaatttaa tttgatcatg ttggcatttt    8880
ttattccttg agtgcattga aaataactga ttaatactag acacatataa tactatcaag    8940
ttacacctaa aatttatttt gtcagccagg caaccaatca gaaccaaata gtgaccaaga    9000
attatttttt caaacacgtt tgtcttccag gaccttactg agttagttcc acagatcatg    9060
aaaacatggg tcttttctca ggtcttctac atcatttccc ttaagttgac acccaaagtc    9120
ataagcttga ctttagacac aggtttcaaa agtagaaagt tttattctct tatgtttcta    9180
tggctttgaa gttttgcaga actataaaaa taatttgttt tacccttttcc acaactcatg    9240
gacgttcctt gaagtgaagg aaacgaatac tcgtctttgc ataataaaca gcaaaaaaaa    9300
aaaaaaaaaa aaaattaaat gaatctaaat tctcatatgc catttaaagt ttcagaccca    9360
gacatcttca ggattttctt acatttaaag cagaatgaca ctactgatac acaaaagtaa    9420
attttaattc tgtttttaat gcatgaggca atataaaaat taatcatgat gatcattata    9480
tttgttctaa cactaaagtt tttcccataa tcccccagat ttttcaagtc tctttttctcc    9540
tcatggcaga tgcccaacca cagagagaaa ggttctataa taaaacttta aaaacactga    9600
cctaaaagag aaatgtttta tcgccatgca atattctaaa ttctggcatt tggtggaaaa    9660
actgtacaca ttttttggcaa tggctttagt gtccactggt ctggcagcta taatgagaat    9720
aataaagtat atggtgatga ctgtttactt accattatga gacatcccca ctgcggtacc    9780
cggcgataca gtcaactgtc tttgaccttt gttactactc tcttccgatg atgatgtcgc    9840
```

```
acttattcta tgctgtctca atgttagagg catatcagtc tccactgaag ccaatctatc   9900
tgtgacggca tctttattca cattatcttg tacaaataat cctgttaaca atgcttttat   9960
atcctgtaaa gaatccattt tcaaaatcat gtcaaggtct tctcgaggaa aaatcagtag  10020
aaatagctgt tccagtcttt ctagccttga ttccacttct gtcagatgtg ccctagtcag  10080
cggagacctt ttggttttgg gagagtagcg acactcccag ttgttcttca gacacttggc  10140
gcacttcggt ttttctttgg agcacttgag ctttttaagt cggcaaatat cgcatgcttg  10200
ttcgatagaa gacagtagct tcatggtggc aagcttaagt ttaaacg               10247


<210> 52
<211> 10514
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
      DNA vector sequence

<400> 52
ctaaattgta agcgttaata ttttgttaaa attcgcgtta aatttttgtt aaatcagctc    60
attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga   120
gatagggttg agtgttgttc cagtttggaa caagagtcca ctattaaaga acgtggactc   180
caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc   240
ctaatcaagt tttttggggt cgaggtgccg taaagcacta atcggaacc ctaaagggag    300
ccccgatt agagcttgac ggggaaagcc ggcgaacgtg gcgagaagg aagggaagaa    360
agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac   420
cacacccgcc gcgcttaatg cgccgctaca gggcgcgtcc cattcgccat tcaggctgcg   480
caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg   540
gggatgtgct gcaaggcgat taagttgggt aacgccaggg tttcccagt cacgacgttg    600
taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaat tggagctcca   660
ccgcggtgcg gccgggccat gcaggccaga catgataaga tacattgatg agtttggaca   720
aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg atgctattgc   780
tttatttgta accattataa gctgcaataa acaagttgtc gagttacttg tacagctcgt   840
ccatgccgag agtgatcccg gcggcggtca cgaactccag caggaccatg tgatcgcgct   900
tctcgttggg gtctttgctc agggcggact ggtagctcag gtagtggttg tcgggcagca   960
gcacggggcc gtcgccgatg ggggtgttct gctggtagtg gtcggcgagc tgcacgctgc  1020
cgtcctcgat gttgtggcgg atcttgaagt tcaccttgat gccgttcttc tgcttgtcgg  1080
ccatgatata gacgttgtgg ctgttgtagt tgtactccag cttgtgcccc aggatgttgc  1140
cgtcctcctt gaagtcgatg cccttcagct cgatgcggtt caccagggtg tcgccctcga  1200
acttcacctc ggcgcgggtc ttgtagttgc cgtcgtcctt gaagaagatg gtgcgctcct  1260
ggacgtagcc ttcgggcatg gcggacttga agaagtcgtg ctgcttcatg tggtcggggt  1320
agcgggcgaa gcactgcagg ccgtagccga aggtggtcac gagggtgggc cagggcacgg  1380
gcagcttgcc ggtggtgcag atgaacttca gggtcagctt gccgtaggtg gcatcgccct  1440
cgccctcgcc ggacacgctg aacttgtggc cgtttacgtc gccgtccagc tcgaccagga  1500
tgggcaccac cccggtgaac agctcctcgc ccttgctcac catggtggcg ttagcttgat  1560
ttgacagtgg ctggggggttg cgccgccggg ttttatagga agccacagcg gccactcgag  1620
ccataaaagg caactttagg aacggcgggg ggtgattgga ttcgagtcgt ttattcaccg  1680
gccttgccgc acagtgcagc atttttttac cccctctccc ctccttttgc gggggaaaaa  1740
aaaaaaaaaa aaaaaaaaag gagaagagaa aaaagcgag cgagagagaa agcgagattg    1800
aggaagagga tgaagagttt ggcgatgggt gctggttccg taggcccaga tggacaagaa  1860
tagcccccgc ccttgcggac agtatcccat tcagtgactc agatcagtca agcggccgcg  1920
tctaacctag gagcccacca cagacactat actataacca cactgaattt tctctcctat  1980
ctcaaataat cctattcttg ttcatgcctc agtgccttat aacaagcaga tttctcggcc  2040
agacatggtg gcttgcgcct gtaatcccag cactttggcc aagtcgggca gctcacttga  2100
ggtcaggagt ttgaattcag cctgtccaat atggtgaaac cccatctcta ctaaaaatac  2160
aaaaattagc tgggcatggt ggtgtgtgcc tgtagtccca gctactcagg agactgaggc  2220
aggagaattg cttgaactcg ggaagcagag gttgcagtga gccaagatca tgccattgca  2280
ctccagcctg ggcaacagag cgagactccg tctcaaacaa aacaaaacaa aaacaaaaaa  2340
```

```
acaagcagat ctctcaactt ggaatgctgc tgcttctcca tgccaactgg aaaacttctt      2400
tttcttttag actcaactca aatatgatcc tccctgaagt ctctccatct ttcctagaca      2460
caattagcaa ccagttcctc tgtgttcaaa cagcacttcg tgtgtatgtg tatgtatgtg      2520
tatttatcat tattattatc acctttgatc actctccctg aaggtaatag ccaggcactt      2580
agtatagagt aggtatcaac aattgcttat tgaaggaata aatgggtgaa tgacaacttt      2640
gcaaagtttc aattttaat ctattgattt gaaagaatta ttatacatta tttaaaatat       2700
cctttctagt tctaatatct tatgtctata gtagttcgac agttacttta aatcagatat      2760
taaatgattt tcattttaca aagaacagac ctgcctttct caatgtcaca agactattgt      2820
agagaatagc agtgaaataa gtacatgtga aaatacttgg caaagtaata gtgaaaatag      2880
agctgatctc acaggaatgt tttgaaaatt aaaacaccgt atgtaaaata ggtaatctaa      2940
tctgtgggca catatttaaa ctctcaataa atattaactt tgattatttta aaaaatacct     3000
ggtagctttt ttgtgaattc tacaagaacc tgtacatgat tggttgccat ttccgtcaaa      3060
atgagaaaat tttcttctgc actgaattct tcttttaact aaattttaag aaaaaaacca      3120
atgattaaaa agtagataag aaaatgataa tataaaacca ttagactaaa gaaaaactag      3180
gtttagatga taaaatttga aaacccatct tactatttct cagatctttg gagggcaatt      3240
accattatca agataggtgg gagacttcat actcttgaat tgtttaatgt tgctgagagt      3300
ttaaaagtct aacttgaatt ctcagctcta atcatttcaa agtaacatat tcacatataa      3360
attatacctc caaggcaata aaaactccag aaaatattaa ctttaaaccc acatattttc      3420
agatattata catacaattt tatttgttat ttcctgaggc atcctctgtt tgttatatga      3480
atccataata aaatgtagaa gagtctgttg atctgggtg agttcagttt tctcctacac       3540
tggcaagaaa ataaaaggtg ttagggaaga tgtgtataaa gccttctttc tatatgtgca      3600
ttcatctgga gggacctaat agcagaaatc ataaaaaaaa aaaggcataa atacagcagg      3660
aactatgagt tctaagagag tgtataattt ctttctctgg gtcctcaggc aggccatatc      3720
acatttcact gaattttcct atctataaaa tatgaataac aaaatattcc acttgtgtat      3780
ttcttaggaa taaatggtag tctggatatg aagactattg agcccattca tattgaaggt      3840
cactgaaagt tcaagctta attataacta atgaaaggtt ttcatatagt atgactacta      3900
tgaatactat attcatatac tattctacat gaattccacc tggtgaagtc atctgagagt      3960
attcatgtgg aggagggatc aacaaactgt ggcccatgga ccaaatccat tggattttat      4020
aaagttttat tggaacacag tcacactcat ttgtttatgt attgtctctg ctactttttgc      4080
actaaaatgg cagagttgag tagctgttgt agataccatg tggccaaagc cctttaaaga      4140
aaaaaagttt gtcaacccct aaaatagaga gtcaacaatc tggacaattc agaccccaaa      4200
catctcttc tccctctaca gcatcttttg aagagtgtga ataattatgt caactctttg       4260
ttttcttata ccaaagtgtt actgatgaag cactatggca aaataatgct ttcatataaa      4320
atacttttc ttaaataaat acaaaaatac gtatttgtaa aactaaatgt ggataaagtt       4380
ttctttgtca ataatgtaat aaatattgag aagggtgtgg tttggacaat catagcttac      4440
tggtttttca gctaatggca aaataaaatt ccatcaaatg tcctgtgcat tatcacaatg      4500
tgtgtgtcta aacaccttgg atgctgttcc cctattttgt cctaagttgc tcaatggatt      4560
tcttgagaaa tttgtaaccc cagtgggtag gactttctc tattcatcca catacgtgta       4620
ttcaccctgt tttactcatt tataaggatt aatccacagc ctctattgag accatgttct      4680
tcccagtttg agactgtctt ctcaaaaaga gggaaaaaat ggtcatcaat actcagttgg      4740
aaaaaatcca ttttatacac caaatcattt ctgctggcaa acactaggta gtaggactgg      4800
aatggagtca gagagaagga aaactttcac ttttttagttt atgcattgct actatttgaa      4860
tgtttacatc aaagcatgta ttaggaatcc attttaatga acaattggac tagaaattca      4920
gctgtgcctg tggctttaat ttgaaagcct cctggccagc tcactatgta cctagaaact      4980
cagttcctgc cagtcttggc agacaccatt gcatattatt accctgcatg actttgttgt      5040
cgaggtcact tgtcgcaagt cacgaccttc actgtcttct ctagacggcg atacagtcaa      5100
ctgtctttga cctttgttac tactctcttc cgatgatgat gtcgcactta ttctatgctg      5160
tctcaatgtt agaggcatat cagtctccac tgaagccaat ctatctgtga cggcatcttt      5220
attcacatta tcttgtacaa ataatcctgt taacaatgct tttatatcct gtaaagaatc      5280
catttttcaaa atcatgtcaa ggtcttctcg aggaaaaatc agtagaaata gctgttccag     5340
tctttctagc cttgattcca cttctgtcag atgtgcccta gtcagcggag accttttggt      5400
tttgggagag tagcgacact cccagttgtt cttcagacac ttggcgcact tcggttttc       5460
tttggagcac ttgagctttt taagtcggca aatatcgcat gcttgttcga tagaagacag      5520
tagcttcatg gtgtcggatc ccgataagcc agtaagcagt gggttctcta gttagccaga      5580
gagctctgct tatatagacc tcccaccgta cacgcctacc gcccatttgc gtcaatgggg      5640
cggagttgtt acgacatttt ggaaagtccc gttgatttg gtgccaaaac aaactcccat       5700
tgacgtcaat ggggtggaga cttggaaatc cccgtgagtc aaaccgctat ccacgcccat      5760
tgatgtactg ccaaaaccgc atcaccatgg taatagcgat gactaatacg tagatgtact      5820
```

```
gccaagtagg aaagtcccat aaggtcatgt actgggcata atgccaggcg ggccatttac  5880
cgtcattgac gtcaataggg ggcgtacttg gcatatgata cacttgatgt actgccaagt  5940
gggcagttta ccgtaaatac tccacccatt gacgtcaatg gaaagtccct attggcgtta  6000
ctatgggaac atacgtcatt attgacgtca atgggcgggg gtcgttgggc ggtcagccag  6060
gcgggccatt taccgtaagt tatgtaacgc ggaactccat atatgggcta tgaactaatg  6120
accccgtaat tgattactat taataactag tcaataatca atgtcaacgc gtatatctgg  6180
cccgtacatc gcgaagcagc gcaaaacgcc taaccctacc cgggctgcag gaattctacc  6240
gggtagggga ggcgcttttc ccaaggcagt ctggagcatg cgctttagca gccccgctgg  6300
cacttggcgc atcacaagtg gcctctggcc tcgcacacat tccacatcca ccggtagcgc  6360
caaccggctc ccttctttgg tggcccgttc gcgccacctt ctactcctcc cctagtcagg  6420
aagttccccc ccgccccgca gctcgcgtcg tgcaggacgt gacaaatgga agtagcacgt  6480
ctcactagtc tcgtgcagat ggacaagcac cgctgagcaa tggaagcggg taggcctttg  6540
gggcagcggc caatagcagc ttggctcctt cgctttctgg gctcagaggc tgggaagggg  6600
tgggtccggg ggcgggctca ggggcgggct caggggcggg gcgggcgcga aggtcctccg  6660
gacccggcat tctgcacgct tcaaaagcgc acgtctgccg cgctgttctc ctcttcctca  6720
tctccgggcc ttcgacctgc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt  6780
ttctgatcga aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat  6840
ctcgtgcttt cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg  6900
ccgatggttt ctacaaagat cgttatgttt atcggcactt tgcatcggcc gcgctcccga  6960
ttccggaagt gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc  7020
gtgcacaggg tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc  7080
cggtcgcgga ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg  7140
gcccattcgg accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga  7200
ttgctgatcc ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg  7260
tcgcgcaggc tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc  7320
tcgtgcacgc ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg  7380
tcattgactg gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct  7440
tctggaggcc gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc  7500
cggagcttgc aggatcgccg cggctccggg cgtatatgct ccgcattggt cttgaccaac  7560
tctatcagag cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg  7620
acgcaatcgt ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg  7680
cggccgtctg gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca  7740
gcactcgtcc gagggcaaag gaatagagta tgcagaaatt gatgatctat taaacaataa  7800
agatgtccac taaaatggaa gtttttcctg tcatactttg ttaagaaggg tgagaacaga  7860
gtacctacat tttgaatgga aggattggag ctacgggggt gggggtgggg tgggattaga  7920
taaatgcctg ctctttactg aaggctcttt actattgctt tatgataatg tttcatagtt  7980
ggatatcata atttaaacaa gcaaaaccaa attaagggcc agctcattcc tcccactcat  8040
gatctataga tctatagatc tctcgtggga tcattgtttt tctcttgatt cccactttgt  8100
ggttctaagt actgtggttt ccaaatgtgt cagtttcata gcctgaagaa cgagatcagc  8160
agcctctgtt ccacatacac ttcattctca gtattgtttt gccaagttct aattccatca  8220
gattttagtg gacatcttta ttgtttaata gatcatcaat ttctgcatac taagcttatc  8280
gataccgtcg acctcgaggg ggggcccggt acccagcttt tgttcccttt agtgagggtt  8340
aattgcgcgc ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct  8400
cacaattcca cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaatg  8460
agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct  8520
gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg  8580
gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg tcgttcggc tgcggcgagc  8640
ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg  8700
aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct  8760
ggcgtttttc cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca  8820
gaggtggcga aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct  8880
cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc  8940
gggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt  9000
tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc  9060
cggtaactat cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc  9120
cactggtaac aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg  9180
gtggcctaac tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc  9240
agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag  9300
```

```
cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga    9360
tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat    9420
tttggtcatg agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag    9480
ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat    9540
cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc    9600
cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat    9660
accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag    9720
ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg    9780
ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc    9840
tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca    9900
acgatcaagg cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg    9960
tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc   10020
actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta   10080
ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc   10140
aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg   10200
ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc   10260
cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc   10320
aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat   10380
actcatactc ttcctttttc aatattattg aagcatttat cagggttatt gtctcatgag   10440
cggatacata tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc   10500
ccgaaaagtg ccac                                                     10514
```

<210> 53
<211> 168
<212> DNA
<213> Homo sapiens

<400> 53
```
cggaggacag tactccgctc ggaggacagt actccgctcg gaggacagta ctccgctcga      60
ggacagtact ccgctcggag gacagtactc cgatccgtcg actctagtcc gctcggagga     120
cagtactccg ctcggaggac agtactccgc tcggaggaca gtactccg                  168
```

<210> 54
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 54
```
gagaacccac tgcttactgg ct                                                22
```

<210> 55
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 55
```
caagatcgcc ctcgcctttg                                                   20
```

```
<210> 56
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 56
atatggtacc gcagtgggga tgtctcataa tgg                                33


<210> 57
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 57
aattggatcc tcaatcagtc catcacctgg                                   30


<210> 58
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 58
cgtctagaga gtgtgtcctg tcagaagaac                                   30


<210> 59
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 59
atgcggccgc actcctgacc tcaggtgatc c                                 31


<210> 60
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 60
cgtctagaga agacagtgaa ggtcgtgac                                    29
```

```
<210> 61
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 61
gcgcggccgc gtctaaccta ggagcccac                                         29


<210> 62
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 62
gctctagagg acgaatgcca agatctga                                          28


<210> 63
<211> 29
<212> DNA
<213> Artificial Sequence

<220>           .
<223> Description of Artificial Sequence: Synthetic primer

<400> 63
caagcggccg ccagtgtgat ggatatctg                                         29


<210> 64
<211> 5218
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic recombinant
      DNA sequence

<400> 64
gccggagcgg tcgagttctg gaccgaccgg ctcgggttct cccgggactt cgtggaggac       60
gacttcgccg gtgtggtccg ggacgacgtg accctgttca tcagcgcggt ccaggaccag      120
gtggtgccgg acaacaccct ggcctgggtg tgggtgcgcg gcctggacga gctgtacgcc      180
gagtggtcgg aggtcgtgtc cacgaacttc cgggacgcct ccgggccggc catgaccgag      240
atcggcgagc agccgtgggg gcgggagttc gccctgcgcg acccggccgg caactgcgtg      300
cacttcgtgg ccgaggagca ggactgacac gtgctacgag atttcgattc caccgccgcc      360
ttctatgaaa ggttgggctt cggaatcgtt ttccgggacg ccggctggat gatcctccag      420
cgcggggatc tcatgctgga gttcttcgcc caccccaact tgtttattgc agcttataat      480
ggttacaaat aaagcaatag catcacaaat ttcacaaata aagcattttt ttcactgcat      540
tctagttgtg gtttgtccaa actcatcaat gtatcttatc atgtctgtat accgtcgacc      600
tctagctaga gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg      660
```

```
ctcacaattc cacacaacat acgagccgga agcataaagt gtaaagcctg gggtgcctaa    720
tgagtgagct aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac    780
ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt    840
gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg ctgcggcga     900
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca    960
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg   1020
ctggcgtttt ccataggct ccgcccccct gacgagcatc acaaaaatcg acgctcaagt   1080
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc   1140
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct   1200
tcgggaagcg tggcgctttc tcaatgctca cgctgtaggt atctcagttc ggtgtaggtc   1260
gttcgctcca agctgggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta   1320
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca   1380
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag   1440
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag   1500
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt   1560
agcggtggtt ttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa   1620
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg   1680
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga   1740
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta   1800
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc   1860
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg   1920
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga   1980
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt   2040
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt   2100
gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc   2160
caacgatcaa ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc   2220
ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca   2280
gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag   2340
tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg   2400
tcaatacggg ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa   2460
cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa   2520
cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga   2580
gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga   2640
atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg   2700
agcggataca tatttgaatg tatttagaaa aataaacaaa tagggttcc gcgcacattt   2760
ccccgaaaag tgccacctga cgtcgacgga tcgggagatc cgatgttttc tgagttactt   2820
ttgtatcccc accccccctc gaggatcgat ccgctcggag gacagtactc cgctcggagg   2880
acagtactcc gctcggagga cagtactccg ctcgaggaca gtactccgct cggaggacag   2940
tactccgatc cgtcgactct agtccgctcg gaggacagta ctccgctcgg aggacagtac   3000
tccgctcgga ggacagtact ccgctcgagg acagtactcg attcgaggac attttgacac   3060
ccccataata tttttccaga attaacagta taaattgcat ctcttgttca agagttccct   3120
atcactctct ttaatcacta ctcacagtaa cctcaactcc tgcccaagct tggtaccacc   3180
atggacccag aaacgctggt gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg   3240
ggttacatcg aactggatct caacagcggt aagatccttg agagttttcg ccccgaagaa   3300
cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt atcccgtatt   3360
gacgccgggc aagagcaact cggtcgccgc atacactatt ctcagaatga cttggttgag   3420
tactcaccag tcacagaaaa gcatcttacg gatggcatga cagtaagaga attatgcagt   3480
gctgccataa ccatgagtga taacactgcg gccaacttac ttctgacaac gatcggagga   3540
ccgaaggagc taaccgcttt tttgcacaac atgggggatc atgtaactcg ccttgatcgt   3600
tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta   3660
gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct agcttcccgg   3720
caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct gcgctcggcc   3780
cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt   3840
atcattgcag cactggggcc agatggtaag ccctcccgta tcgtagttat ctacacgacg   3900
gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctcactg   3960
attaagcatt ggtaatctag agagggcccg tttaaacccg ctgatcagcc tcgactgtgc   4020
cttctagttg ccagccatct gttgtttgcc cctcccccgt gccttccttg accctggaag   4080
gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcatcgcat tgtctgagta   4140
```

```
ggtgtcattc tattctgggg ggtggggtgg ggcaggacag caagggggag gattgggaag    4200
acaatagcag gcatgctggg gatgcggtgg gctctatggc ttctgaggcg gaaagaacca    4260
gctgggctc  tagggggtat ccccacgcgc cctgtagcgg cgcattaagc gcggcgggtg    4320
tggtggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgccc gctcctttcg    4380
ctttcttccc ttcctttctc gccacgttcg ccggctttcc ccgtcaagct ctaaatcggg    4440
gcatcccttt agggttccga tttagtgctt tacggcacct cgaccccaaa aaacttgatt    4500
agggtgatgg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc cctttgacgt    4560
tggagtccac gttctttaat agtggactct tgttccaaac tggaacaaca ctcaaccta     4620
tctcggtcta ttcttttgat ttataaggga ttttggggat ttcggcctat tggttaaaaa    4680
atgagctgat ttaacaaaaa tttaacgcga attaattctg tggaatgtgt gtcagttagg    4740
gtgtggaaag tccccaggct ccccaggcag gcagaagtat gcaaagcatg catctcaatt    4800
agtcagcaac caggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca    4860
tgcatctcaa ttagtcagca accatagtcc cgcccctaac tccgcccatc ccgcccctaa    4920
ctccgcccag ttccgcccat tctccgcccc atggctgact aatttttttt atttatgcag    4980
aggccgaggc cgcctctgcc tctgagctat tccagaagta gtgaggaggc ttttttggag    5040
gcctaggctt ttgcaaaaag ctcccgggag cttgtatatc cattttcgga tctgatcagc    5100
acgtgttgac aattaatcat cggcatagta tatcggcata gtaataacg  acaaggtgag    5160
gaactaaacc atggccaagt tgaccagtgc cgttccggtg ctcaccgcgc gcgacgtc     5218
```


<210> 65
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 65
ggggtaccaa agaaggtagg ctgagggg   28

<210> 66
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 66
cgggatccgt acaagacagt tagctagttg gc   32

<210> 67
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 67
ggggtaccat tcagcaggcc actacaggac tctc   34

<210> 68
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 68
gatcgcggcc gctgtgctcg acattggtgg cc   32

<210> 69
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 69
ggggtacctt tgtggtgctt aaaaaatgag c   31

<210> 70
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 70
gatcgcggcc gcctcatgaa caatgaaatc tcc   33

<210> 71
<211> 2379
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      Nurr1 targeting sequence

<400> 71
cgtacaagac agttagctag ttggcaaaac caaggagaat tctgtgacaa gggaaactca 60
ggacaaatcc tgctaggcag ttctggaggg gaagcgccct gcgcctgcag tactgacctg 120
tgaccatagc cagggcagca atgcaggaga aggcagaaat gtcgatgttc atattctgca 180
agttggagga gaattcaaca atggaatcaa tccattcccc aaagccacga acgcattgca 240
acctgtgcaa gaccacccca ttgcaaaaga tgagtttacc ctccactggg ttggacctgc 300
aattaatacc aaagagagag aggggagaaa aagagagaga gaaaagctaa acaactaatt 360
acttgaagag caataaatga gggatttaag agtcacctaa ttactgaaga gttaattaaa 420
atgtagatcc agtggacctt gaaagggttt aatttcataa caatcaagaa cgcccccctat 480
cccacctcca ttccattcat ccaaggcttc tgggctcgct tcttctcgtc ttttttttgtt 540

```
ttctttctt  ttcttcctt  ctccgactt  catttcctat  tctgtctttt  tctctacccc  600
accctctggt  ttcccttcct  ccctttcttt  tcctttcttg  atttctctca  cagcctccct  660
ggattgtctc  ctccctccct  tattacctgt  atgctaatcg  aaggacaaac  agttctaaga  720
aagctgattc  aaaaagcagg  tcttggtcgg  ctttgggcag  gtctgcgaag  ccagggatct  780
tctctgccca  gccccggatg  atctccatgg  agccagtcag  gagatcatag  aattgctgga  840
tatgctgggt  gtcatctcca  ctcatttgat  agtcagggtt  cgcctggaac  tggaatttca  900
ttttaaaaag  cacttaatga  ggttctctaa  aatatataac  ccgtgaaatt  gctaaccccg  960
tttctaatag  gggagccagg  tttttataac  aattaaacct  ctctctgacc  agtaaggaaa  1020
ttagatgttc  cggggcaatt  aattcaatta  gggatggtgc  tacgaggtcg  ctgctttaaa  1080
tatgtaaatt  gtcatttcca  tacagactaa  atacagtgcc  atccagccct  gggaaacgtt  1140
cactcttatc  tccacaaaac  aattgacatg  ttagtattca  tgctagtccc  agagtgggcc  1200
tcgagctgag  acgccctttt  gcaaatctta  ataaaattgc  agtccctacc  aggtttgctc  1260
aacatacttc  ctccaattaa  agcccgcata  tttttcacag  ctggagtaaa  aggatctgac  1320
cacttgatcc  cccaccccac  tggcatcttt  acaccctctc  cttcccttcc  ctgaccccca  1380
gacagccccc  acatccttcc  ccagcacagt  gccaggaagg  atgaggatta  aagcatctat  1440
ctctatggag  gactgggaat  aaagtagatt  gggaagggg  agagtttgca  cagcccctgg  1500
atttcatggt  agaggcactg  gtgtatatgg  aagaggaggg  gcgatggggt  agtggggtag  1560
tgggaattca  cacagcttca  aaccctgtgg  cgtgtccctc  tcctccttcc  tccatcctgt  1620
cctcatgtct  atgcccccag  ctacatggat  ttgtctgaag  atctgaatct  tgaggccctg  1680
gccagagctg  cgaggcatat  acagccttgc  ttgcctttct  ttacccccgt  tgaatctgag  1740
agttaatgac  ggatgtgggg  aggggtcctg  cccatctgtc  ggtttgtcca  catgatatcc  1800
cccccgccag  cttcttaccc  tggaatagtc  caggctggtc  atagccgggt  tggagtcgac  1860
atgggccctg  acgagggcac  tgatcagact  caccgggggc  gaagggggag  agggctcctg  1920
tgggctcttc  ggtttcgagg  gcaaacgacc  tctccggcct  tttaaactgt  ctgtgcgaac  1980
cactgcaaag  gaagagccct  gttagcgccg  cttttccgag  cccaggccca  gctgctgcct  2040
cggtccctcc  ccggggaagg  ccgcagccgc  ggggcaccag  gctgagcggc  tgagggcccc  2100
agtgcttgta  aagccttcac  tgactagaag  cattaaaaaa  tgcggggtta  ttttatgtct  2160
tcctccaaat  gggtcgtata  gttaaaggag  agaagggcct  ggcggctttc  tctagggaag  2220
gccgggcaag  caggcagctg  cagggtcctg  gaggccatac  tgaggggag   tcggagatcc  2280
ccagcaccgg  gaagtggaac  gtgatgctgg  agtatgagca  gtggtttcct  aaaggcgcaa  2340
actggagggt  cggcagctcc  cctcagccta  ccttctttg                           2379
```

<210> 72
<211> 8426
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      pCDGal4-DBD-Nurr1 sequence

<400> 72
```
tctctggcta  actagagaac  ccactgctta  ctggcttatc  gaaattaata  cgactcacta  60
tagggagacc  caagctggct  agcgtttaaa  cttaagcttg  ccaccatgaa  gctactgtct  120
tctatcgaac  aagcatgcga  tatttgccga  cttaaaaagc  tcaagtgctc  caaagaaaaa  180
ccgaagtgcg  ccaagtgtct  gaagaacaac  tgggagtgtc  gctactctcc  caaaaccaaa  240
aggtctccgc  tgactagggc  acatctgaca  gaagtggaat  caaggctaga  aagactggaa  300
cagctatttc  tactgatttt  tcctcgagaa  gaccttgaca  tgattttgaa  aatggattct  360
ttacaggata  taaaagcatt  gttaacagga  ttatttgtac  aagataatgt  gaataaagat  420
gccgtcacag  atagattggc  ttcagtggag  actgatatgc  ctctaacatt  gagacagcat  480
agaataagtg  cgacatcatc  atcggaagag  agtagtaaca  aaggtcaaag  acagttgact  540
gtatcgccgg  gtaccaaaga  aggtaggctg  aggggagctg  ccgaccctcc  agtttgcgcc  600
tttaggaaac  cactgctcat  actccagcat  cacgttccac  ttcccggtgc  tggggatctc  660
cgactccccc  tcagtatggc  ctccaggacc  ctgcagctgc  ctgcttgccc  ggccttccct  720
agagaaagcc  gccaggccct  tctctccttt  aactatcga   cccatttgga  ggaagacata  780
aaataacccc  gcatttttta  atgcttctag  tcagtgaagg  ctttacaagc  actggggccc  840
tcagccgctc  agcctggtgc  cccgcggctg  cggccttccc  cggggaggga  ccgaggcagc  900
```

```
agctgggcct gggctcggaa aagcggcgct aacagggctc ttcctttgca gtggttcgca 960
cagacagttt aaaaggccgg agaggtcgtt tgccctcgaa accgaagagc ccacaggagc 1020
cctctccccc ttcgccccg gtgagtctga tcagtgccct cgtcagggcc catgtcgact 1080
ccaacccggc tatgaccagc ctggactatt ccagggtaag aagctggcgg gggggatatc 1140
atgtggacaa accgacagat gggcaggacc cctccccaca tccgtcatta actctcagat 1200
tcaacggggg taaagaaagg caagcaaggc tgtatatgcc tcgcagctct ggccagggcc 1260
tcaagattca gatcttcaga caaatccatg tagctggggg catagacatg aggacaggat 1320
ggaggaagga ggagagggac acgccacagg gtttgaagct gtgtgaattc ccactacccc 1380
actaccccat cgcccctcct cttccatata caccagtgcc tctaccatga aatccagggg 1440
ctgtgcaaac tctcccccctt cccaatctac tttattccca gtcctccata gagatagatg 1500
ctttaatcct catccttcct ggcactgtgc tggggaagga tgtgggggct gtctgggggt 1560
cagggaaggg aaggagaggg tgtaaagatg ccagtggggt gggggatcaa gtggtcagat 1620
ccttttactc cagctgtgaa aaatatgcgg gctttaattg gaggaagtat gttgagcaaa 1680
cctggtaggg actgcaattt tattaagatt tgcaaaaggg cgtctcagct cgaggcccac 1740
tctgggacta gcatgaatac taacatgtca attgttttgt ggagataaga gtgaacgttt 1800
cccagggctg gatggcactg tatttagtct gtatggaaat gacaatttac atatttaaag 1860
cagcgacctc gtagcaccat ccctaattga attaattgcc ccggaacatc taatttcctt 1920
actggtcaga gagaggttta attgttataa aaacctggct cccctattag aaacgggggtt 1980
agcaatttca cgggttatat attttagaga acctcattaa gtgctttta aaatgaaatt 2040
ccagttccag gcgaaccctg actatcaaat gagtggagat gacacccagc atatccagca 2100
attctatgat ctcctgactg gctccatgga gatcatccgg ggctgggcag agaagatccc 2160
tggcttcgca gacctgccca aagccgacca agacctgctt tttgaatcag ctttcttaga 2220
actgtttgtc cttcgattag catacaggta ataagggagg gaggagacaa tccagggagg 2280
ctgtgagaga aatcaagaaa ggaaaagaaa gggaggaagg gaaaccagag ggtggggtag 2340
agaaaaagac agaataggaa atggaagtcg gagaaaggaa gaaaaagaaa gaaaacaaaa 2400
aaagacgaga agaagcgagc ccagaagcct tggatgaatg gaatggaggt gggatagggg 2460
gcgttcttga ttgttatgaa attaaaccct ttcaaggtcc actggatcta cattttaatt 2520
aactcttcag taattaggtg actcttaaat ccctcattta ttgctcttca agtaattagt 2580
tgtttagctt ttctctctct ctttttctcc cctctcctc tttggtatta attgcaggtc 2640
caacccagtg gagggtaaac tcatcttttg caatggggtg gtcttgcaca ggttgcaatg 2700
cgttcgtggc tttggggaat ggattgattc cattgttgaa ttctcctcca acttgcagaa 2760
tatgaacatc gacatttctg ccttctcctg cattgctgcc ctggctatgg tcacaggtca 2820
gtactgcagg cgcagggcgc ttcccctcca gaactgccta gcaggatttg tcctgagttt 2880
cccttgtcac agaattctcc ttggttttgc caactagcta actgtcttgt acggatccac 2940
tagtaacggc cgccagtgtg ctggaattct gcagatatcc atcacactgg cggccgctcg 3000
agtctagagg gcccgtttaa acccgctgat cagcctcgac tgtgccttct agttgccagc 3060
catctgttgt ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc actcccactg 3120
tcctttccta ataaaatgag gaaattgcat cgcattgtct gagtaggtgt cattctattc 3180
tggggggtgg ggtggggcag gacagcaagg gggaggattg ggaagacaat agcaggcatg 3240
ctggggatgc ggtgggctct atggcttctg aggcggaaag aaccagctgg ggctctaggg 3300
ggtatcccca cgcgccctgt agcggcgcat taagcgcggc gggtgtggtg gttacgcgca 3360
gcgtgaccgc tacacttgcc agcgccctag cgcccgctcc tttcgctttc ttcccttcct 3420
ttctcgccac gttcgccggc tttccccgtc aagctctaaa tcggggggctc cctttagggt 3480
tccgatttag tgctttacgg cacctcgacc ccaaaaaact tgattagggt gatggttcac 3540
gtagtgggcc atcgccctga tagacggttt ttcgcccttt gacgttggag tccacgttct 3600
ttaatagtgg actcttgttc caaactggaa caacactcaa ccctatctcg gtctattctt 3660
ttgatttata agggattttg ccgatttcgg cctattggtt aaaaaatgag ctgatttaac 3720
aaaaatttaa cgcgaattaa ttctgtggaa tgtgtgtcag ttagggtgtg aaagtcccc 3780
aggctcccca gcaggcagaa gtatgcaaag catgcatctc aattagtcag caaccaggtg 3840
tggaaagtcc ccaggctccc cagcaggcag aagtatgcaa agcatgcatc tcaattagtc 3900
agcaaccata gtcccgcccc taactccgcc catcccgccc ctaactccgc ccagttccgc 3960
ccattctccg ccccatggct gactaatttt ttttatttat gcagaggccg aggccgcctc 4020
tgcctctgag ctattccaga agtagtgagg aggcttttt ggaggcctag gcttttgcaa 4080
aaagctcccg ggagcttgta tatccatttt cggatctgat cagcacgtga tgaaaaagcc 4140
tgaactcacc gcgacgtctg tcgagaagtt tctgatcgaa aagttcgaca gcgtctccga 4200
cctgatgcag ctctcggagg gcgaagaatc tcgtgctttc agcttcgatg taggagggcg 4260
tggatatgtc ctgcgggtaa atagctgcgc cgatggtttc tacaaagatc gttatgttta 4320
tcggcacttt gcatcggccg cgctcccgat tccggaagtg cttgacattg gggaattcag 4380
```

```
cgagagcctg acctattgca tctcccgccg tgcacagggt gtcacgttgc aagacctgcc 4440
tgaaaccgaa ctgcccgctg ttctgcagcc ggtcgcggag gccatggatg cgatcgctgc 4500
ggccgatctt agccagacga gcgggttcgg cccattcgga ccgcaaggaa tcggtcaata 4560
cactacatgg cgtgatttca tatgcgcgat tgctgatccc catgtgtatc actggcaaac 4620
tgtgatggac gacaccgtca gtgcgtccgt cgcgcaggct ctcgatgagc tgatgctttg 4680
ggccgaggac tgccccgaag tccggcacct cgtgcacgcg gatttcggct ccaacaatgt 4740
cctgacggac aatggccgca taacagcggt cattgactgg agcgaggcga tgttcggggga 4800
ttcccaatac gaggtcgcca acatcttctt ctggaggccg tggttggctt gtatggagca 4860
gcagacgcgc tacttcgagc ggaggcatcc ggagcttgca ggatcgccgc ggctccgggc 4920
gtatatgctc cgcattggtc ttgaccaact ctatcagagc ttggttgacg gcaatttcga 4980
tgatgcagct tgggcgcagg gtcgatgcga cgcaatcgtc cgatccggag ccgggactgt 5040
cgggcgtaca caaatcgccc gcagaagcgc ggccgtctgg accgatggct gtgtagaagt 5100
actcgccgat agtggaaacc gacgccccag cactcgtccg agggcaaagg aatagcacgt 5160
gctacgagat ttcgattcca ccgccgcctt ctatgaaagg ttgggcttcg gaatcgtttt 5220
ccgggacgcc ggctggatga tcctccagcg cggggatctc atgctggagt tcttcgccca 5280
ccccaacttg tttattgcag cttataatgg ttacaaataa agcaatagca tcacaaattt 5340
cacaaataaa gcattttttt cactgcattc tagttgtggt ttgtccaaac tcatcaatgt 5400
atcttatcat gtctgtatac cgtcgacctc tagctagagc ttggcgtaat catggtcata 5460
gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag 5520
cataaagtgt aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg 5580
ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca 5640
acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc 5700
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 5760
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 5820
ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gcccccctga 5880
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga acccgacag gactataaag 5940
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 6000
taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 6060
ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 6120
ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 6180
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattacga gcgaggta 6240
tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 6300
agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 6360
ttgatccggc aaacaaacca ccgctggtag cggttttttt gtttgcaagc agcagattac 6420
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca 6480
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 6540
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 6600
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 6660
tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt 6720
accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt 6780
atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc 6840
cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa 6900
tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg 6960
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat cccccatgtt 7020
gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc 7080
agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt 7140
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 7200
gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 7260
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 7320
gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt cagcatcttt 7380
tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg 7440
aataagggcg acacggaaat gttgaatact catactcttc cttttttcaat attattgaag 7500
catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 7560
acaaataggg gttccgcgca catttccccg aaaagtgcca cctgacgtcg acggatcggg 7620
agatctcccg atcccctatg gtgcactctc agtacaatct gctctgatgc cgcatagtta 7680
agccagtatc tgctccctgc ttgtgtgttg gaggtcgctg agtagtgcgc gagcaaaatt 7740
taagctacaa caaggcaagg cttgaccgac aattgcatga agaatctgct tagggttagg 7800
cgttttgcgc tgcttcgcga tgtacgggcc agatatacgc gttgacattg attattgact 7860
```

```
agttattaat agtaatcaat tacggggtca ttagttcata gcccatatat ggagttccgc 7920
gttacataac ttacggtaaa tggcccgcct ggctgaccgc ccaacgaccc ccgcccattg 7980
acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca ttgacgtcaa 8040
tgggtggagt atttacggta aactgcccac ttggcagtac atcaagtgta tcatatgcca 8100
agtacgcccc ctattgacgt caatgacggt aaatggcccg cctggcatta tgcccagtac 8160
atgaccttat gggactttcc tacttggcag tacatctacg tattagtcat cgctattacc 8220
atggtgatgc ggttttggca gtacatcaat gggcgtggat agcggtttga ctcacgggga 8280
tttccaagtc tccacccat tgacgtcaat gggagtttgt tttggcacca aaatcaacgg 8340
gactttccaa aatgtcgtaa caactccgcc ccattgacgc aaatgggcgg taggcgtgta 8400
cggtgggagg tctatataag cagagc                                    8426
```

<210> 73
<211> 9701
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      pKI-Gal4-DBD-Nurr1 sequence

<220>
<221> modified_base
<222> (1518)..(1518)
<223> any nucleotide

<400> 73
```
tgcacgcttc aaaagcgcac gtctgccgcg ctgttctcct cttcctcatc tccgggcctt 60
cgacctgcat gaaaaagcct gaactcaccg cgacgtctgt cgagaagttt ctgatcgaaa 120
agttcgacag cgtctccgac ctgatgcagc tctcggaggg cgaagaatct cgtgctttca 180
gcttcgatgt aggagggcgt ggatatgtcc tgcgggtaaa tagctgcgcc gatggtttct 240
acaaagatcg ttatgtttat cggcactttg catcggccgc gctcccgatt ccggaagtgc 300
ttgacattgg ggaattcagc gagagcctga cctattgcat ctcccgccgt gcacagggtg 360
tcacgttgca agacctgcct gaaaccgaac tgcccgctgt tctgcagccg gtcgcggagg 420
ccatggatgc gatcgctgcg gccgatctta gccagacgag cgggttcggc ccattcggac 480
cgcaaggaat cggtcaatac actacatggc gtgatttcat atgcgcgatt gctgatcccc 540
atgtgtatca ctggcaaact gtgatggacg acaccgtcag tgcgtccgtc gcgcaggctc 600
tcgatgagct gatgctttgg ccgaggact gccccgaagt ccggcacctc gtgcacgcgg 660
atttcggctc caacaatgtc ctgacggaca atggccgcat aacagcggtc attgactgga 720
gcgaggcgat gttcggggat tcccaatacg aggtcgccaa catcttcttc tggaggccgt 780
ggttggcttg tatggagcag cagacgcgct acttcgagcg gaggcatccg gagcttgcag 840
gatcgccgcg gctccgggcg tatatgctcc gcattggtct tgaccaactc tatcagagct 900
tggttgacgg caatttcgat gatgcagctt gggcgcaggg tcgatgcgac gcaatcgtcc 960
gatccggagc cgggactgtc gggcgtacac aaatcgcccg cagaagcgcg gccgtctgga 1020
ccgatggctg tgtagaagta ctcgccgata gtggaaaccg acgccccagc actcgtccga 1080
gggcaaagga atgcctgaga aaggaagtga gctgtaaagg ctgagctctc tctctgacgt 1140
atgtagcctc tggttagctt cgtcactcac tgttcttgac tcagcatggc aatctgatga 1200
aatcccagct gtaagtctgc ataaattgat gatctattaa acaataaaga tgtccactaa 1260
aatggaagtt ttttactgtc atactttgtt aagaagggtg agaacagagt acctacattt 1320
tgaatggaag gattggagct acggggtgg gggtggggt gggattagat aaatgcctgc 1380
tctttactga aggctcttta ctattgcttt atgataatgt ttcatagttg gatatcataa 1440
tttaaacaag caaaaccaaa ttaagggcca gctcattcct ccactcatga tctatagatc 1500
tatagatctc tcgtgggatc attgtttntc tgatccactg gaagcttatc gataccgtcg 1560
acctcgaggg ggggcccggt acccagcttt tgttcccttt agtgagggtt aattgcgcgc 1620
ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct cacaattcca 1680
cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaatg agtgagctaa 1740
ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct gtcgtgccag 1800
ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc 1860
```

```
gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct 1920
cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 1980
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 2040
cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 2100
aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 2160
cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg 2220
gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag 2280
ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat 2340
cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac 2400
aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac 2460
tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc agttaccttc 2520
ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt 2580
tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc 2640
ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg 2700
agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca 2760
atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca 2820
cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag 2880
ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac 2940
ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc 3000
agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct 3060
agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc 3120
gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg 3180
cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc 3240
gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat 3300
tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag 3360
tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat 3420
aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg 3480
cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca 3540
cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga 3600
aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc 3660
ttcctttttc aatattattg aagcatttat cagggttatt gtctcatgag cggatacata 3720
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 3780
ccacctaaat tgtaagcgtt aatattttgt taaaattcgc gttaaatttt tgttaaatca 3840
gctcattttt taaccaatag gccgaaatcg gcaaaatccc ttataaatca aaagaataga 3900
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 3960
actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 4020
caccctaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 4080
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 4140
agaaagcgaa aggagcgggc gctagggcgc tggcaagtgt agcggtcacg ctgcgcgtaa 4200
ccaccacacc cgccgcgctt aatgcgccgc tacagggcgc gtcccattcg ccattcaggc 4260
tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc gctattacgc cagctggcga 4320
aagggggatg tgctgcaagg cgattaagtt gggtaacgcc agggttttcc cagtcacgac 4380
gttgtaaaac gacggccagt gagcgcgcgt aatacgactc actatagggc gaattggagc 4440
tccaccgcgg tgcggccggg ccatgcaggc cacgacatga taagatacat tgatgagttt 4500
ggacaaacca caactagaat gcagtgaaaa aaatgcttta tttgtgaaat ttgtgatgct 4560
attgctttat ttgtaaccat tataagctgc aataaacaag ttcttgtaca gctcgtccat 4620
gccgagagtg atcccggcgg cggtcacgaa ctccagcagg accatgtgat cgcgcttctc 4680
gttggggtct ttgctcaggg cggactggta gctcaggtag tggttgtcgg gcagcagcac 4740
ggggccgtcg ccgatggggg tgttctgctg gtagtggtcg gcgagctgca cgctgccgtc 4800
ctcgatgttg tggcggatct tgaagttcac cttgatgccg ttcttctgct tgtcggccat 4860
gatatagacg ttgtggctgt tgtagttgta ctccagcttg tgccccagga tgttgccgtc 4920
ctccttgaag tcgatgccct tcagctcgat gcggttcacc agggtgtcgc cctcgaactt 4980
cacctcggcg cgggtcttgt agttgccgtc gtccttgaag aagatggtgc gctcctggac 5040
gtagccttcg ggcatggcgg acttgaagaa gtcgtgctgc ttcatgtggt cggggtagcg 5100
ggcgaagcac tgcaggccgt agccgaaggt ggtcacgagg gtgggccagg gcacgggcag 5160
cttgccggtg gtgcagatga acttcagggt cagcttgccg taggtggcat cgccctcgcc 5220
ctcgccggac acgctgaact tgtggccgtt tacgtcgccg tccagctcga ccaggatggg 5280
caccaccccg gtgaacagct cctcgccctt gctcaccatg gtggcgttag cttgatttga 5340
```

```
cagtggctgg gggttgcgcc gccgggtttt ataggaagcc acagcggcca ctcgagccat 5400
aaaaggcaac tttaggaacg gcggggggtg attggattcg agtcgtttat tcaccggcct 5460
tgccgcacag tgcagcattt ttttacccac tctcccctcc ttttgcgggg gaaaaaaaaa 5520
aaaaaaaaaa aaaaggaga agagaaaaaa agcgagcgag agagaaagcg agattgagga 5580
agaggatgaa gagtttggcg atgggtgctg gttccgtagg cccagatgga caagaatagc 5640
ccccgccctt gcggacagta tcccattcag tgactcagat cagatcaagc ggccgccagt 5700
gtgatggata tctgcagaat tccagcacac tggcggccgt tactagtgga tccgtacaag 5760
acagttagct agttggcaaa accaaggaga attctgtgac aagggaaact caggacaaat 5820
cctgctaggc agttctggag gggaagcgcc ctgcgcctgc agtactgacc tgtgaccata 5880
gccagggcag caatgcagga gaaggcagaa atgtcgatgt tcatattctg caagttggag 5940
gagaattcaa caatggaatc aatccattcc ccaaagccac gaacgcattg caacctgtgc 6000
aagaccaccc cattgcaaaa gatgagttta ccctccactg ggttggacct gcaattaata 6060
ccaaagagag agagggggaga aaaagagaga gagaaaagct aaacaactaa ttacttgaag 6120
agcaataaat gagggattta agagtcacct aattactgaa gagttaatta aaatgtagat 6180
ccagtggacc ttgaaagggt ttaatttcat aacaatcaag aacgccccct atcccacctc 6240
cattccattc atccaaggct tctgggctcg cttcttctcg tcttttttg ttttcttct 6300
ttttcttcct ttctccgact tccatttcct attctgtctt tttctctacc ccaccctctg 6360
gtttcccttc ctcccttct tttcctttct tgatttctct cacagcctcc ctggattgtc 6420
tcctccctcc cttattacct gtatgctaat cgaaggacaa acagttctaa gaaagctgat 6480
tcaaaaagca ggtcttggtc ggctttgggc aggtctgcga agccagggat cttctctgcc 6540
cagccccgga tgatctccat ggagccagtc aggagatcat agaattgctg gatatgctgg 6600
gtgtcatctc cactcatttg atagtcaggg ttcgcctgga actggaattt cattttaaaa 6660
agcacttaat gaggttctct aaaatatata acccgtgaaa ttgctaaccc cgtttctaat 6720
aggggagcca ggttttata acaattaaac ctctctctga ccagtaagga aattagatgt 6780
tccggggcaa ttaattcaat tagggatggt gctacgaggt cgctgcttta aatatgtaaa 6840
ttgtcatttc catacagact aaatacagtg ccatccagcc ctgggaaacg ttcactctta 6900
tctccacaaa acaattgaca tgttagtatt catgctagtc ccagagtggg cctcgagctg 6960
agacgccctt ttgcaaatct taataaaatt gcagtcccta ccaggtttgc tcaacatact 7020
tcctccaatt aaagcccgca tattttcac agctggagta aaaggatctg accacttgat 7080
ccccccacccc actggcatct ttacaccctc tccttccctt ccctgacccc cagacagccc 7140
ccacatcctt ccccagcaca gtgccaggaa ggatgaggat taaagcatct atctctatgg 7200
aggactggga ataaagtaga ttgggaaggg ggagagtttg cacagcccct ggatttcatg 7260
gtagaggcac tggtgtatat ggaagaggag gggcgatggg gtagtggggt agtgggaatt 7320
cacacagctt caaaccctgt ggcgtgtccc tctcctcctt cctccatcct gtcctcatgt 7380
ctatgccccc agctacatgg atttgtctga agatctgaat cttgaggccc tggccagagc 7440
tgcgaggcat atacagcctt gcttgccttt cttaccccc gttgaatctg agagttaatg 7500
acggatgtgg ggaggggtcc tgcccatctg tcggtttgtc cacatgatat ccccccgcc 7560
agcttcttac cctggaatag tccaggctgg tcatagccgg gttggagtcg acatgggccc 7620
tgacgagggc actgatcaga ctcaccgggg gcgaaggggg agagggctcc tgtgggctct 7680
tcggtttcga gggcaaacga cctctccggc ctttaaact gtctgtgcga accactgcaa 7740
aggaagagcc ctgttagcgc cgctttccg agcccaggcc cagctgctgc ctcggtccct 7800
ccccggggaa ggccgcagcc gcggggcacc aggctgagcg gctgagggcc ccagtgcttg 7860
taaagccttc actgactaga agcattaaaa aatgcggggt tattttatgt cttcctccaa 7920
atgggtcgta tagttaaagg agagaagggc ctggcggctt tctctaggga aggccgggca 7980
agcaggcagc tgcaggtcc tggaggccat actgaggggg agtcggagat ccccagcacc 8040
gggaagtgga acgtgatgct ggagtatgag cagtggtttc ctaaaggcgc aaactggagg 8100
gtcggcagct cccctcagcc taccttcttt ggtacccggc gatacagtca actgtctttg 8160
acctttgtta ctactctctt ccgatgatga tgtcgcactt attctatgct gtctcaatgt 8220
tagaggcata tcagtctcca ctgaagccaa tctatctgtg acggcatctt tattcacatt 8280
atcttgtaca aataatcctg ttaacaatgc ttttatatcc tgtaaagaat ccattttcaa 8340
aatcatgtca aggtcttctc gaggaaaaat cagtagaaat agctgttcca gtctttctag 8400
ccttgattcc acttcgtca gatgtgccct agtcagcgga gaccttttgg ttttgggaga 8460
gtagcgacac tcccagttgt tcttcagaca cttggcgcac ttcggttttt ctttggagca 8520
cttgagcttt ttaagtcggc aaatatcgca tgcttgttcg atagaagaca gtagcttcat 8580
ggtggcaagc ttaagtttaa acgctagaac tagtggatcc ggataagcca gtaagcagtg 8640
ggttctctag ttagccagag agctctgctt atatagacct cccaccgtac acgcctaccg 8700
cccatttgcg tcaatggggc ggagttgtta cgacattttg gaaagtcccg ttgattttgg 8760
tgccaaaaca aactcccatt gacgtcaatg gggtggagac ttggaaatcc ccgtgagtca 8820
```

```
aaccgctatc cacgcccatt gatgtactgc caaaaccgca tcaccatggt aatagcgatg 8880
actaatacgt agatgtactg ccaagtagga aagtcccata aggtcatgta ctgggcataa 8940
tgccaggcgg gccatttacc gtcattgacg tcaatagggg gcgtacttgg catatgatac 9000
acttgatgta ctgccaagtg ggcagtttac cgtaaatact ccacccattg acgtcaatgg 9060
aaagtcccta ttggcgttac tatgggaaca tacgtcatta ttgacgtcaa tgggcggggg 9120
tcgttgggcg gtcagccagg cgggccattt accgtaagtt atgtaacgcg gaactccata 9180
tatgggctat gaactaatga ccccgtaatt gattactatt aataactagt caataatcaa 9240
tgtcaacccc gggctgcagg aattctaccg ggtaggggag gcgcttttcc caaggcagtc 9300
tggagcatgc gctttagcag ccccgctggc acttggcgca tcacaagtgg cctctggcct 9360
cgcacacatt ccacatccac cggtagcgcc aaccggctcc cttctttggt ggcccccttcg 9420
cgccaccttc tactcctccc ctagtcagga agttcccccc cgccccgcag ctcgcgtcgt 9480
gcaggacgtg acaaatggaa gtagcacgtc tcactagtct cgtgcagatg gacaagcacc 9540
gctgagcaat ggaagcgggt aggcctttgg ggcagcggcc aatagcagct tggctccttc 9600
gctttctggg ctcagaggct gggaaggggt gggtccgggg gcgggctcag gggcgggctc 9660
aggggcgggg cgggcgcgaa ggtcctccgg acccggcatt c 9701
```

<210> 74
<211> 2379
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic nucleotide
      sequence for GR region of homology

<400> 74
```
cgtacaagac agttagctag ttggcaaaac caaggagaat tctgtgacaa gggaaactca 60
ggacaaatcc tgctaggcag ttctggaggg gaagcgccct gcgcctgcag tactgacctg 120
tgaccatagc cagggcagca atgcaggaga aggcagaaat gtcgatgttc atattctgca 180
agttggagga gaattcaaca atggaatcaa tccattcccc aaagccacga acgcattgca 240
acctgtgcaa gaccacccca ttgcaaaaga tgagtttacc ctccactggg ttggacctgc 300
aattaatacc aaagagagag aggggagaaa aagagagaga gaaaagctaa acaactaatt 360
acttgaagag caataaatga gggatttaag agtcacctaa ttactgaaga gttaattaaa 420
atgtagatcc agtggacctt gaaagggttt aatttcataa caatcaagaa cgcccccat 480
cccacctcca ttccattcat ccaaggcttc tgggctcgct tcttctcgtc ttttttgtt 540
ttctttcttt ttcttccttt ctccgacttc catttcctat tctgtctttt tctctacccc 600
accctctggt ttcccttcct ccctttcttt tcctttcttg atttctctca cagcctccct 660
ggattgtctc ctccctccct tattacctgt atgctaatcg aaggacaaac agttctaaga 720
aagctgattc aaaaagcagg tcttggtcgg ctttgggcag gtctgcgaag ccagggatct 780
tctctgccca gccccggatg atctccatgg agccagtcag gagatcatag aattgctgga 840
tatgctgggt gtcatctcca ctcatttgat agtcagggtt cgcctggaac tggaatttca 900
ttttaaaaag cacttaatga ggttctctaa aatatataac ccgtgaaatt gctaaccccg 960
tttctaatag gggagccagg tttttataac aattaaacct ctctctgacc agtaaggaaa 1020
ttagatgttc cggggcaatt aattcaatta gggatggtgc tacgaggtcg ctgctttaaa 1080
tatgtaaatt gtcatttcca tacagactaa atacagtgcc atccagccct gggaaacgtt 1140
cactcttatc tccacaaaac aattgacatg ttagtattca tgctagtccc agagtgggcc 1200
tcgagctgag acgccctttt gcaaatctta ataaaattgc agtccctacc aggtttgctc 1260
aacatacttc ctccaattaa agcccgcata tttttcacag ctggagtaaa aggatctgac 1320
cacttgatcc cccaccccac tggcatcttt acaccctctc cttcccttcc ctgaccccca 1380
gacagccccc acatccttcc ccagcacagt gccaggaagg atgaggatta aagcatctat 1440
ctctatggag gactgggaat aaagtagatt gggaaggggg agagtttgca cagcccctgg 1500
atttcatggt agaggcactg gtgtatatgg aagaggaggg gcgatggggt agtggggtag 1560
tgggaattca cacagcttca aaccctgtgg cgtgtccctc tcctccttcc tccatcctgt 1620
cctcatgtct atgcccccag ctacatggat ttgtctgaag atctgaatct tgaggccctg 1680
gccagagctg cgaggcatat acagccttgc ttgcctttct ttaccccgt tgaatctgag 1740
agttaatgac ggatgtgggg aggggtcctg cccatctgtc ggtttgtcca catgatatcc 1800
cccccgccag cttcttaccc tggaatagtc caggctggtc atagccgggt tggagtcgac 1860
```

```
atgggccctg acgagggcac tgatcagact caccggggggc gaaggggggag agggctcctg 1920
tgggctcttc ggtttcgagg gcaaacgacc tctccggcct tttaaactgt ctgtgcgaac 1980
cactgcaaag gaagagccct gttagcgccg cttttccgag cccaggccca gctgctgcct 2040
cggtccctcc ccggggaagg ccgcagccgc ggggcaccag gctgagcggc tgagggcccc 2100
agtgcttgta aagccttcac tgactagaag cattaaaaaa tgcggggtta ttttatgtct 2160
tcctccaaat gggtcgtata gttaaaggag agaagggcct ggcggctttc tctagggaag 2220
gccgggcaag caggcagctg cagggtcctg gaggccatac tgaggggggag tcggagatcc 2280
ccagcaccgg gaagtggaac gtgatgctgg agtatgagca gtggtttcct aaaggcgcaa 2340
actggagggt cggcagctcc cctcagccta ccttctttg 2379
```

<210> 75
<211> 9597
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
pCDGal4-DBD-GR sequence

<400> 75
```
taacaggatt atttgtacaa gataatgtga ataaagatgc cgtcacagat agattggctt 60
cagtggagac tgatatgcct ctaacattga gacagcatag aataagtgcg acatcatcat 120
cggaagagag tagtaacaaa ggtcaaagac agttgactgt atcgccgggt accattcagc 180
aggccactac aggagtctca caagaaacct ctgaaaatcc tggtaacaaa acaatagttc 240
ctgcaacgtt accacaactc accccctaccc tggtgtcact gttggaggtt attgaacctg 300
aagtgttata tgcaggatat gatagctctg ttccagactc aacttggagg atcatgacta 360
cgctcaacat gttaggaggg cggcaagtga ttgcagcagt gaaatgggca aaggcaatac 420
caggtaagat gcaaaacata aaagagcaac tatataaacc tttgtgtttt cttcagcaaa 480
aacactttgg cttttatatc atcgtgagcc catggcttat cttgtttctc ttagttctgg 540
ggactatgaa ggggagagtc aggtgaatac aggtgatagg gagtttataa taaaacattt 600
acattactcc ctgcttttca aatcattatg cacaggatgg taatttcaca taggatgatg 660
taatatcaga attcaagtta caagactcac tcaaaactcc ttttacactg aagtttgggg 720
aaagaaaatg ttttagtta attccatttg ttttccttca ttgtgccact tttaaaaatc 780
aggttgtttg taagattggt aaacatcaag tatgttgatt gtcaaaattt gtactaaagt 840
agaatgattt taacccttca ctaaatgaaa tgctacacat tgaatgtaat tttaaagata 900
atttaaata aaagttaccc tattggaatt tggtgtggaa tggcagaggt caatgttagt 960
gtcagctctg actttaaaga cagggaattg acaagcctgt gttcacgcaa atagttaggg 1020
agagagcaag aaagtaacct gacctcctgt catccttgtt ttattaaggg ggaaagaggt 1080
gtgaatagca gggcaaatgt tttgcttaac tcattgatta atacctcaag ccaagattct 1140
tttctgtttt ttaaaatcaa tacataatag ttgtacatat ttactgtaca tatttatatt 1200
taggggggtac atgtaataat ttaataaaag catacaacgt gtaaggatca aatcagagta 1260
actgggatat ccatcacctc aaacatttgt ttggggaaca ttccaaatct tctctttag 1320
ctattttgaa atataaagta aattattgtt aactatagtc atcctgttgt gctactgaac 1380
actaaaactt atttcttcta actgtatttt tgcacccgtc aaccattccc gcttcatccc 1440
catcaccact atctttcccg gtcactggta accgccaagc caagaatttt ggctatttta 1500
ctatttagtt catgtttact taagcagaca gaggtgacaa aactggcttt ttttttttt 1560
ttttacatta aaagctatta aaaagcacct aggggggctgg gtgcgatggc tcacgcctgt 1620
aatcccagca ctttgggaag cccaggtggg tggatcagtt gaggtcagga gttcgagacc 1680
agcctggcca gcatagcaaa accccatctc tactaaaatt acaaaaatta gccgggcatg 1740
gtggtatgaa tctgtattcc tagctacttg ggaggctggc actgagaatc acttgaaccc 1800
gggaggcgga ggttgcagtg agccgagatg gcaccattgc actccagcct gggtgacaga 1860
gcaagacttt gtctcaatta aaaaaaaaaa aaaaaaaaaa aacacaagag ggtttgtgag 1920
tcttaaagtg tcagatgaca gaagaaaact gtgtctacct agtatttaat ttccattttc 1980
tgttagggggt gcccttgttt tgacagggct aattgatctc attgctcctt ggcaattccc 2040
acagagatga tcttctgaag agtgttgcct cataccttta tttctcttaa ttcaggtttc 2100
aggaacttac acctggatga ccaaatgacc ctactgcagt actcctggat gtttcttatg 2160
gcatttgctc tggggtggag atcatataga caatcaagtg caaacctgct gtgttttgct 2220
```

```
cctgatctga ttattaatga gtaagttgta tgtgtgtcat tttccctgta ttcatagggt 2280
atctttaacc agctgatgtt ttcctgattg actgctattg tgataattca ggactgaaac 2340
aatcctacta ggtatctagg atctaggcaa actggaaata gagttatgag tgcttggggc 2400
aggacaagtg taatgtaaag caaatgtaca tgtggcatta ttactgtccc aggacatgtt 2460
tgaggatatt taacagcata tctgaggtta gtaaagtctg tcgcaagcaa caaggaatct 2520
tactgtgata tcatttacat aaccctattc cagaaagaaa aaggagcatg gtaaaactca 2580
tgtggattca gtggggacaa ttgtagatga ggatatctag gctgatgggg tgggacatat 2640
ggacccagac acaagaggta tctctttgca tggcaaggct cacccagtgt ctgtggttta 2700
agaatatggg aacaaatttg ttttgtttaa ctgagagaag accaagcctt taagatttta 2760
taaatcagct attctcttat cctctaagct tattcctgtg tctgcgaaat acttcaggtg 2820
tccatttccc cttacctcat tgcagttgtt tcctcactcg ttttctccct ccagtgtaac 2880
gttcatcatg ttggctaatg tttgcttcct caagcacagt ctgactgcat cacatatctc 2940
cccagtacac agattgtctt cagtatcttc ccactgaccc tccagtacat attctgcatg 3000
atttcagact ttccagaatc tgacctcact tcctctccca ttgtttttcct tcacacactc 3060
ttcattccca tccatccttt ccagcatact cttagactct tggtgttcac atcaccagat 3120
acacagcaga gaagtcacat cctagttact ctcactttct accttgtatt actacttttc 3180
gtacccctag cttattgcta ttagtacaat gtaaacaggg agttcacaca cacataccc 3240
tggtctaaga agaataaaaa atgaaggaga tttctgtttg tatagaaaac agaagtcacc 3300
ttgactttta ttgccaaaaa gaggactgtt caaactactg catcacaatg taacaagatt 3360
aggtagttgg atccaatttt aaattaactg gtaaatatat ttagtttctg gggaaactga 3420
agacattatt actcatcata atcctaccat gctgtttaaa aaataccatg ttggcagtat 3480
ttgttttta gtcactttct aatatgtaat ttgaaggcat ttaagtggaa ttaaaagcat 3540
aaacagattt gtatgaaaca ccaacttatc ctggtttata aaactaacct aatttagggt 3600
ttttattatt agggcattca gatttagctt taagcagtca cagcaaaatc taatcatgcc 3660
acatacattc cttacataaa gtgggattta taattttttt tcctcaacag atttacatta 3720
gtttcatttt cattaaggga tatgtacttc ctattcttgt gttctcatgc tgctgcctgc 3780
ggccgctcga gtctagaggg cccgtttaaa cccgctgatc agcctcgact gtgccttcta 3840
gttgccagcc atctgttgtt tgcccctccc ccgtgccttc cttgaccctg gaaggtgcca 3900
ctcccactgt cctttcctaa taaaatgagg aaattgcatc gcattgtctg agtaggtgtc 3960
attctattct ggggggtggg gtggggcagg acagcaaggg ggaggattgg gaagacaata 4020
gcaggcatgc tggggatgcg gtggggctcta tggcttctga ggcggaaaga accagctggg 4080
gctctagggg gtatccccac gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg 4140
ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct ttcgctttct 4200
tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat cggggggctcc 4260
ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt gattagggtg 4320
atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggagt 4380
ccacgttctt taatagtgga ctcttgttcc aaactggaac aacactcaac cctatctcgg 4440
tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc 4500
tgatttaaca aaaatttaac gcgaattaat tctgtggaat gtgtgtcagt tagggtgtgg 4560
aaagtcccca ggctccccag caggcagaag tatgcaaagc atgcatctca attagtcagc 4620
aaccaggtgt ggaaagtccc caggctcccc agcaggcaga agtatgcaaa gcatgcatct 4680
caattagtca gcaaccatag tcccgcccct aactccgccc atcccgcccc taactccgcc 4740
cagttccgcc cattctccgc cccatggctg actaattttt tttatttatg cagaggccga 4800
ggccgcctct gcctctgagc tattccagaa gtagtgagga ggcttttttg gaggcctagg 4860
cttttgcaaa aagctcccgg gagcttgtat atccattttc ggatctgatc agcacgtgat 4920
gaaaaagcct gaactcaccg cgacgtctgt cgagaagttt ctgatcgaaa agttcgacag 4980
cgtctccgac ctgatgcagc tctcggaggg cgaagaatct cgtgctttca gcttcgatgt 5040
aggagggcgt ggatatgtcc tgcgggtaaa tagctgcgcc gatggtttct acaaagatcg 5100
ttatgtttat cggcactttg catcggccgc gctcccgatt ccggaagtgc ttgacattgg 5160
ggaattcagc gagagcctga cctattgcat ctcccgccgt gcacagggtg tcacgttgca 5220
agacctgcct gaaaccgaac tgcccgctgt tctgcagccg tcgcggagg ccatggatgc 5280
gatcgctgcg gccgatctta ccagacgag cgggttcggc ccattcggac cgcaaggaat 5340
cggtcaatac actacatggc gtgatttcat atgcgcgatt gctgatcccc atgtgtatca 5400
ctggcaaact gtgatggacg acaccgtcag tgcgtccgtc gcgcaggctc tcgatgagct 5460
gatgctttgg gccgaggact gccccgaagt ccggcacctc gtgcacgcgg atttcggctc 5520
caacaatgtc ctgacggaca atggccgcat aacagcggtc attgactgga gcgaggcgat 5580
gttcggggat cccaatacg aggtcgccaa catcttcttc tggaggccgt ggttggcttg 5640
tatggagcag cagacgcgct acttcgagcg gaggcatccg gagcttgcag atcgccgcg 5700
```

```
gctccgggcg tatatgctcc gcattggtct tgaccaactc tatcagagct tggttgacgg 5760
caatttcgat gatgcagctt gggcgcaggg tcgatgcgac gcaatcgtcc gatccggagc 5820
cgggactgtc gggcgtacac aaatcgcccg cagaagcgcg gccgtctgga ccgatggctg 5880
tgtagaagta ctcgccgata gtggaaaccg acgccccagc actcgtccga gggcaaagga 5940
atagcacgtg ctacgagatt tcgattccac cgccgccttc tatgaaaggt tgggcttcgg 6000
aatcgttttc cgggacgccg gctggatgat cctccagcgc ggggatctca tgctggagtt 6060
cttcgcccac cccaacttgt ttattgcagc ttataatggt tacaaataaa gcaatagcat 6120
cacaaatttc acaaataaag cattttttc actgcattct agttgtggtt tgtccaaact 6180
catcaatgta tcttatcatg tctgtatacc gtcgacctct agctagagct tggcgtaatc 6240
atggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac acaacatacg 6300
agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac tcacattaat 6360
tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc tgcattaatg 6420
aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg cttcctcgct 6480
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 6540
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt gagcaaaagg 6600
ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg 6660
cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg 6720
actataaaga taccaggcgt ttccccctgg aagctccctc gtgcgctctc ctgttccgac 6780
cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca 6840
tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt 6900
gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc 6960
caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca ggattagcag 7020
agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac 7080
tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg gaaaaagagt 7140
tggtagctct tgatccggca aacaaaccac cgctggtagc ggtttttttg tttgcaagca 7200
gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc 7260
tgacgctcag tggaacgaaa actcacgtta agggatttg gtcatgagat tatcaaaaag 7320
gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata 7380
tgagtaaact tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat 7440
ctgtctattt cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg 7500
ggagggctta ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc 7560
tccagattta tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc 7620
aactttatcc gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc 7680
gccagttaat agtttgcgca acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc 7740
gtcgtttggt atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc 7800
ccccatgttg tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa 7860
gttggccgca gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat 7920
gccatccgta agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata 7980
gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca 8040
tagcagaact ttaaaagtgc tcatcattgg aaaacgttct cggggcgaa aactctcaag 8100
gatcttaccg ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc 8160
agcatctttt actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc 8220
aaaaaaggga ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata 8280
ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta 8340
gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcga 8400
cggatcggga gatctcccga tcccctatgg tgcactctca gtacaatctg ctctgatgcc 8460
gcatagttaa gccagtatct gctccctgct tgtgtgttgg aggtcgctga gtagtgcgcg 8520
agcaaaattt aagctacaac aaggcaaggc ttgaccgaca attgcatgaa gaatctgctt 8580
agggttaggc gttttgcgct gcttcgcgat gtacgggcca gatatacgcg ttgacattga 8640
ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg 8700
gagttccgcg ttacataact tacggtaaat ggcccgcctg gctgaccgcc caacgacccc 8760
cgcccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat 8820
tgacgtcaat gggtggagta tttacggtaa actgcccact ggcagtacaa tcaagtgtat 8880
catatgccaa gtacgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat 8940
gcccagtaca tgaccttatg ggactttcct acttggcagt acatctacgt attagtcatc 9000
gctattacca tggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac 9060
tcacggggat ttccaagtct ccacccccat tgacgtcaatg ggagtttgtt ttggcaccaa 9120
aatcaacggg actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt 9180
```

```
aggcgtgtac ggtgggaggt ctatataagc agagctctct ggctaactag agaacccact 9240
gcttactggc ttatcgaaat taatacgact cactataggg agacccaagc tggctagcgt 9300
ttaaacttaa gcttgccacc atgaagctac tgtcttctat cgaacaagca tgcgatattt 9360
gccgacttaa aaagctcaag tgctccaaag aaaaaccgaa gtgcgccaag tgtctgaaga 9420
acaactggga gtgtcgctac tctcccaaaa ccaaaaggtc tccgctgact agggcacatc 9480
tgacagaagt ggaatcaagg ctagaaagac tggaacagct atttctactg attttttcctc 9540
gagaagacct tgacatgatt ttgaaaatgg attctttaca ggatataaaa gcattgt    9597
```

```
<210> 76
<211> 10872
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      pKI-Gal4-DBD-GR sequence

<220>
<221> modified_base
<222> (8885)..(8885)
<223> any nucleotide

<400> 76
cagttcgatg taacccactc gtgcacccaa ctgatcttca gcatctttta ctttcaccag 60
cgtttctggg tgagcaaaaa caggaaggca aaatgccgca aaaaagggaa taagggcgac 120
acggaaatgt tgaatactca tactcttcct ttttcaatat tattgaagca tttatcaggg 180
ttattgtctc atgagcggat acatatttga atgtatttag aaaaataaac aaataggggt 240
tccgcgcaca tttccccgaa aagtgccacc taaattgtaa gcgttaatat tttgttaaaa 300
ttcgcgttaa attttttgtta aatcagctca ttttttaacc aataggccga atatcggcaaa 360
atcccttata aatcaaaaga atagaccgag ataggttga gtgttgttcc agtttggaac 420
aagagtccac tattaaagaa cgtggactcc aacgtcaaag ggcgaaaaac cgtctatcag 480
ggcgatggcc cactacgtga accatcaccc taatcaagtt ttttggggtc gaggtgccgt 540
aaagcactaa atcggaaccc taaagggagc ccccgattta gagcttgacg gggaaagccg 600
gcgaacgtgg cgagaaagga agggaagaaa gcgaaaggag cgggcgctag ggcgctggca 660
agtgtagcgg tcacgctgcg cgtaaccacc acacccgccg cgcttaatgc gccgctacag 720
ggcgcgtccc attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc 780
tcttcgctat tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta 840
acgccagggt tttcccagtc acgacgttgt aaaacgacgg ccagtgagcg cgcgtaatac 900
gactcactat agggcgaatt ggagctccac cgcggtgcgg ccgggccatg caggccacga 960
catgataaga tacattgatg agtttggaca aaccacaact agaatgcagt gaaaaaaatg 1020
ctttatttgt gaaatttgtg atgctattgc tttatttgta accattataa gctgcaataa 1080
acaagttctt gtacagctcg tccatgccga gagtgatccc ggcggcggtc acgaactcca 1140
gcaggaccat gtgatcgcgc ttctcgttgg ggtctttgct cagggcggac tggtagctca 1200
ggtagtggtt gtcgggcagc agcacggggc cgtcgccgat ggggtgttc tgctggtagt 1260
ggtcggcgag ctgcacgctg ccgtcctcga tgttgtggcg gatcttgaag ttcaccttga 1320
tgccgttctt ctgcttgtcg gccatgatat agacgttgtg gctgttgtag ttgtactcca 1380
gcttgtgccc caggatgttg ccgtcctcct tgaagtcgat gcccttcagc tcgatgcggt 1440
tcaccagggt gtcgccctcg aacttcacct cggcgcgggt cttgtagttg ccgtcgtcct 1500
tgaagaagat ggtgcgctcc tggacgtagc cttcgggcat ggcggacttg aagaagtcgt 1560
gctgcttcat gtggtcgggg tagcgggcga agcactgcag ccgtagccg aaggtggtca 1620
cgagggtggg ccagggcacg ggcagcttgc cggtggtgca gatgaacttc agggtcagct 1680
tgccgtaggt ggcatcgccc tcgccctcgc cggacacgct gaacttgtgg ccgtttacgt 1740
cgccgtccag ctcgaccagg atgggcacca ccccggtgaa cagctcctcg cccttgctca 1800
ccatggtggc gttagcttga tttgacagtg ctgggggtt gcgccgccgg gttttatagg 1860
aagccacagc ggccactcga gccataaaag gcaactttag gaacggcggg gggtgattgg 1920
attcgagtcg tttattcacc ggccttgccg cacagtgcag cattttttta cccctctcc 1980
cctcctttg cggggaaaa aaaaaaaaaa aaaaaaaaaa ggagaagaga aaaaagcga 2040
```

```
gcgagagaga aagcgagatt gaggaagagg atgaagagtt tggcgatggg tgctggttcc 2100
gtaggcccag atggacaaga atagcccccg cccttgcgga cagtatccca ttcagtgact 2160
cagatcagat caagcggccg caggcagcag catgagaaca caagaatagg aagtacatat 2220
cccttaatga aaatgaaact aatgtaaatc tgttgaggaa aaaaaattat aaatcccact 2280
ttatgtaagg aatgtatgtg gcatgattag attttgctgt gactgcttaa agctaaatct 2340
gaatgcccta ataataaaaa ccctaaatta ggttagtttt ataaaccagg ataagttggt 2400
gtttcataca aatctgttta tgctttaat tccacttaaa tgccttcaaa ttacatatta 2460
gaaagtgact aaaaaacaaa tactgccaac atggtatttt ttaaacagca tggtaggatt 2520
atgatgagta ataatgtctt cagtttcccc agaaactaaa tatatttacc agttaattta 2580
aaattggatc caactaccta atcttgttac attgtgatgc agtagtttga acagtcctct 2640
ttttggcaat aaaagtcaag gtgacttctg ttttctatac aaacagaaat ctccttcatt 2700
ttttattctt cttagaccag gggtatgtgt gtgtgaactc cctgtttaca ttgtactaat 2760
agcaataagc tagggggtacg aaaagtagta atacaaggta gaaagtgaga gtaactagga 2820
tgtgacttct ctgctgtgta tctggtgatg tgaacaccaa gagtctaaga gtatgctgga 2880
aaggatggat gggaatgaag agtgtgtgaa ggaaaacaat gggagaggaa gtgaggtcag 2940
attctggaaa gtctgaaatc atgcagaata tgtactggag ggtcagtggg aagatactga 3000
agacaatctg tgtactgggg agatatgtga tgcagtcaga ctgtgcttga ggaagcaaac 3060
attagccaac atgatgaacg ttacactgga gggagaaaac gagtgaggaa acaactgcaa 3120
tgaggtaagg ggaaatggac acctgaagta tttcgcagac acaggaataa gcttagagga 3180
taagagaata gctgatttat aaaatcttaa aggcttggtc ttctctcagt taaacaaaac 3240
aaatttgttc ccatattctt aaaccacaga cactgggtga gccttgccat gcaaagagat 3300
acctcttgtg tctgggtcca tatgtcccac cccatcagcc tagatatcct catctacaat 3360
tgtccccact gaatccacat gagtttacc atgctccttt ttctttctgg aatagggtta 3420
tgtaaatgat atcacagtaa gattccttgt tgcttgcgac agactttact aacctcagat 3480
atgctgttaa atatcctcaa acatgtcctg ggacagtaat aatgccacat gtacatttgc 3540
tttacattac acttgtcctg ccccaagcac tcataactct atttccagtt tgcctagatc 3600
ctagatacct agtaggattg tttcagtcct gaattatcac aatagcagtc aatcaggaaa 3660
acatcagctg gttaaagata ccctatgaat acagggaaaa tgacacacat acaacttact 3720
cattaataat cagatcagga gcaaaacaca gcaggtttgc acttgattgt ctatatgatc 3780
tccacccag agcaaatgcc ataagaaaca tccaggagta ctgcagtagg gtcatttggt 3840
catccaggtg taagttcctg aaacctgaat taagagaaat aaaggtatga ggcaacactc 3900
ttcagaagat catctctgtg ggaattgcca aggagcaatg agatcaatta gccctgtcaa 3960
aacaagggca cccctaacag aaaatggaaa ttaaatacta ggtagacaca gtttcttct 4020
gtcatctgac actttaagac tcacaaaccc tcttgtgttt tttttttttt tttttttttt 4080
aattgagaca aagtcttgct ctgtcaccca ggctggagtg caatggtgcc atctcggctc 4140
actgcaacct ccgcctcccg ggttcaagtg attctcagtg ccagcctccc aagtagctag 4200
gaatacagat tcataccacc atgcccggct aatttttgta attttagtag agatggggtt 4260
ttgctatgct ggccaggctg gtctcgaact cctgacctca actgatccac ccacctgggc 4320
ttcccaaagt gctgggatta caggcgtgag ccatcgcacc cagcccccta ggtgcttttt 4380
aatagctttt aatgtaaaaa aaaaaaaaaa aagccagttt tgtcacctct gtctgcttaa 4440
gtaaacatga actaaatagt aaaatagcca aaattcttgg cttggcggtt accagtgacc 4500
gggaaagata gtggtgatgg ggatgaagcg ggaatggttg acgggtgcaa aaatacagtt 4560
agaagaaata agtttagtg ttcagtagca caacaggatg actatagtta acaataattt 4620
actttatatt tcaaaatagc taaaagagaa gatttggaat gttccccaaa caaatgtttg 4680
aggtgatgga tatcccagtt actctgattt gatccttaca cgttgtatgc ttttattaaa 4740
ttattacatg tacccccctaa atataaatat gtacagtaaa tatgtacaac tattatgtat 4800
tgattttaaa aaacagaaaa gaatcttggc ttgaggtatt aatcaatgag ttaagcaaaa 4860
catttgccct gctattcaca cctctttccc ccttaataaa acaaggatga caggaggtca 4920
ggttactttc ttgctctctc cctaactatt tgcgtgaaca caggcttgtc aattccctgt 4980
ctttaaagtc agagctgaca ctaacattga cctctgccat tccacaccaa attccaatag 5040
ggtaactttt atttaaaatt atctttaaaa ttacattcaa tgtgtagcat ttcatttagt 5100
gaagggttaa aatcattcta ctttagtaca aattttgaca atcaacatac ttgatgttta 5160
ccaatcttac aaacaacctg atttttaaaa gtggcacaat gaaggaaaac aaatggaatt 5220
aactaaaaac attttctttc cccaaacttc agtgtaaaag gagttttgag tgagtcttgt 5280
aacttgaatt ctgatattac atcatcctat gtgaaattac catcctgtcg ataatgattt 5340
gaaaagcagg gagtaatgta aatgtttttat tataaactcc ctatcacctg tattcacctg 5400
actctcccct tcatagtccc cagaactaag agaaacaaga taagccatgg gctcacgatg 5460
atataaaagc caaagtgttt ttgctgaaga aaacacaaag gtttatatag ttgctctttt 5520
```

```
atgttttgca tcttacctgg tattgccttt gcccatttca ctgctgcaat cacttgccgc 5580
cctcctaaca tgttgagcgt agtcatgatc ctccaagttg agtctggaac agagctatca 5640
tatcctgcat ataacacttc aggttcaata acctccaaca gtgacaccag ggtaggggtg 5700
agttgtggta acgttgcagg aactattgtt ttgttaccag gattttcaga ggtttcttgt 5760
gagactcctg tagtggcctg ctgaatggta cccggcgata cagtcaactg tctttgacct 5820
ttgttactac tctcttccga tgatgatgtc gcacttattc tatgctgtct caatgttaga 5880
ggcatatcag tctcccactga agccaatcta tctgtgacgg catctttatt cacattatct 5940
tgtacaaata atcctgttaa caatgctttt atatcctgta aagaatccat tttcaaaatc 6000
atgtcaaggt cttctcgagg aaaaatcagt agaaatagct gttccagtct ttctagcctt 6060
gattccactt ctgtcagatg tgccctagtc agcggagacc ttttggtttt gggagagtag 6120
cgacactccc agttgttctt cagacacttg gcgcacttcg gttttctttt ggagcacttg 6180
agcttttttaa gtcggcaaat atcgcatgct tgttcgatag aagacagtag cttcatggtg 6240
gcaagcttaa gtttaaacgc tagaactagt ggatccggat aagccagtaa gcagtgggtt 6300
ctctagttag ccagagagct ctgcttatat agacctccca ccgtacacgc ctaccgccca 6360
tttgcgtcaa tggggcggag ttgttacgac attttggaaa gtcccgttga ttttggtgcc 6420
aaaacaaact cccattgacg tcaatggggt ggagacttgg aaatccccgt gagtcaaacc 6480
gctatccacg cccattgatg tactgccaaa accgcatcac catggtaata gcgatgacta 6540
atacgtagat gtactgccaa gtaggaaagt cccataaggt catgtactgg gcataatgcc 6600
aggcgggcca tttaccgtca ttgacgtcaa taggggggcgt acttggcata tgatacactt 6660
gatgtactgc caagtgggca gtttaccgta aatactccac ccattgacgt caatggaaag 6720
tccctattgg cgttactatg ggaacatacg tcattattga cgtcaatggg cgggggtcgt 6780
tgggcggtca gccaggcggg ccatttaccg taagttatgt aacgcggaac tccatatatg 6840
ggctatgaac taatgacccc gtaattgatt actattaata actagtcaat aatcaatgtc 6900
aaccccgggc tgcaggaatt ctaccgggta ggggaggcgc ttttcccaag gcagtctgga 6960
gcatgcgctt tagcagcccc gctggcactt ggcgcatcac aagtggcctc tggcctcgca 7020
cacattccac atccaccggt agcgccaacc ggctcccttc tttggtggcc ccttcgcgcc 7080
accttctact cctcccctag tcaggaagtt ccccccccgcc ccgcagctcg cgtcgtgcag 7140
gacgtgacaa atggaagtag cacgtctcac tagtctcgtg cagatggaca agcaccgctg 7200
agcaatggaa gcgggtaggc ctttggggca gcggccaata gcagcttggc tccttcgctt 7260
tctgggctca gaggctggga aggggtgggt ccggggggcgg gctcaggggc gggctcaggg 7320
gcggggcggg cgcgaaggtc ctccggaccc ggcattctgc acgcttcaaa agcgcacgtc 7380
tgccgcgctg ttctcctctt cctcatctcc gggccttcga cctgcatgaa aaagcctgaa 7440
ctcaccgcga cgtctgtcga gaagtttctg atcgaaaagt tcgacagcgt ctccgacctg 7500
atgcagctct cggagggcga agaatctcgt gctttcagct tcgatgtagg agggcgtgga 7560
tatgtcctgc gggtaaatag ctgcgccgat ggtttctaca aagatcgtta tgtttatcgg 7620
cactttgcat cggccgcgct cccgattccg gaagtgcttg acattgggga attcagcgag 7680
agcctgacct attgcatctc ccgccgtgca cagggtgtca cgttgcaaga cctgcctgaa 7740
accgaactgc ccgctgttct gcagccggtc gcggaggcca tggatgcgat cgctgcggcc 7800
gatcttagcc agacgagcgg gttcggccca ttcggaccgc aaggaatcgg tcaatacact 7860
acatggcgtg atttcatatg cgcgattgct gatccccatg tgtatcactg gcaaactgtg 7920
atggacgaca ccgtcagtgc gtccgtcgcg caggctctcg atgagctgat gctttgggcc 7980
gaggactgcc ccgaagtccg gcacctcgtg cacgcggatt cggctccaa caatgtcctg 8040
acggacaatg gccgcataac agcggtcatt gactggagcg aggcgatgtt cggggattcc 8100
caatacgagg tcgccaacat cttcttctgg aggccgtggt tggcttgtat ggagcagcag 8160
acgcgctact tcgagcggag gcatccggag cttgcaggat cgccgcggct ccgggcgtat 8220
atgctccgca ttggtcttga ccaactctat cagagcttgg ttgacggcaa tttcgatgat 8280
gcagcttggg cgcagggtcg atgcgacgca atcgtccgat ccggagccgg gactgtcggg 8340
cgtacacaaa tcgcccgcag aagcgcggcc gtctggaccg atggctgtgt agaagtactc 8400
gccgatagtg gaaaccgacg ccccagcact cgtccgaggg caaaggaatg cctgagaaag 8460
gaagtgagct gtaaaggctg agctctctct ctgacgtatg tagcctctgg ttagcttcgt 8520
cactcactgt tcttgactca gcatggcaat ctgatgaaat cccagctgta agtctgcata 8580
aattgatgat ctattaaaca ataaagatgt ccactaaaat ggaagttttt tactgtcata 8640
ctttgttaag aagggtgaga acagagtacc tacattttga atggaaggat tggagctacg 8700
ggggtggggg tgggggtggg attagataaa tgcctgctct ttactgaagg ctctttacta 8760
ttgctttatg ataatgtttc atagttggat atcataattt aaacaagcaa aaccaaatta 8820
agggccagct cattcctcca ctcatgatct atagatctat agatctctcg tgggatcatt 8880
gtttntctga tccactggaa gcttatcgat accgtcgacc tcgaggggggg gcccggtacc 8940
cagcttttgt tccctttagt gagggttaat tgcgcgcttg gcgtaatcat ggtcatagct 9000
```

```
gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 9060
aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgctc 9120
actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg 9180
cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct 9240
gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt 9300
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 9360
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 9420
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 9480
ccaggcgttt cccCctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac 9540
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 9600
taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 9660
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag 9720
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 9780
aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaaggacagt 9840
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 9900
atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 9960
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca 10020
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 10080
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 10140
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 10200
tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt 10260
accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt 10320
atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc 10380
cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa 10440
tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg 10500
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat cccccatgtt 10560
gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc 10620
agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt 10680
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 10740
gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 10800
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 10860
gctgttgaga tc                                                     10872
```

<210> 77
<211> 3250
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic nucleotide
      sequence for MR region of homology sequence

<400> 77
```
tttgtggtgc ttaaaaaatg agcatgtggt gcctcctatg gtccaggtta ataccagact 60
ttgataaagt tctatggtag aacatgtttc cctggcacct gtcctccctg cctgacagac 120
tcaagctctt cttttgttct ttattttatt tatttatttt ttaagatgga ttctcattct 180
gtcgcccagg ctggagttca gtgacatgat attggctcac tgcatcccag gtttcaagcg 240
attctcctgc ctcaggctca caagtagctg gggttacagg tgcctgccag gacgcccaac 300
taaattttgt atttttagta gagacagggt ttcatgatgg tggcaaggct ggtctcaaat 360
tcctgacctc aagtgatctg cttgcctcag cctcccaaag tctgggatta caggcatggg 420
ccactgcacc aggccttctt ttgttttta tcacctcccc catctccagg ggcatcatgt 480
tatttttctg cctggagt ttcactgggc aagctcctct tcaaagaatc agtcagtgga 540
gctttcccca acttttacag ctgctgtctg accacaggta atctaagaaa ctgctttaga 600
ttaatgaact gggacaatgt aatggcaatt tctaagtgat gtttctgtag attaagaggg 660
gccttattaa ttaaatacca tgaaacaaaa aaaaatgact aaaataggaa ttccagggac 720
aatttaaagg gtgctggggg ccaggtgatt attttttaga cagaagagtc attccaggtc 780
tcttacaact tacttgctag gccttgcctt acttggtcta cacttagatt aaccctcgtt 840
```

```
atgagccgct aataatgtgc tcaacttgtg atttcatgaa cagaatgtat tttaccttat 900
taatatttt ggatgaaatt ttccattttg atcaacataa gaagtgtatg tcatttttt 960
tatttttaat ttcttttttt ttttgagac agagtctcac tctgttgctc aggctggagt 1020
atagtggtgg gatctcggct cactgcagcc tctgccgccc aggttcaagc gattctcctg 1080
cctcagcctc ctgagtagct gggattacag gcacctgcca ccaagcagga ctaattttg 1140
tgtttttagt agaaacgggg tttcaccatc ttggctaggc tggtcttgaa ctcctgacct 1200
cgtgatccat ctgcctcagt aagccacctg ccttggcctc ccaaagtgct gggattacag 1260
gcgtgagcca ctgcgcccag ccgtatctgt cattttaaca ataaaatcct taaccctgca 1320
ttctcggaag aaatgagcac attaaccata tttttcctat acacttgatc aatttgttat 1380
caatttaaaa aatctgatat caggatttca tagaacgcaa actcctcctg catggttgga 1440
gatggcgaag tcagttgccc agacttgtta agttcaggat gcagcctgtg aaaggagagg 1500
caatcctacc tggaagtacc tttgcccact tcacgacttg gatcatctgt ttgcctgcta 1560
agcggttgag cgtggagagc agattttcgg ctgtatctgg ttttgagctg tcatagcctg 1620
catatacaat ttcaggttca atgtttttcaa ggaccataac ggggggaaggt gtgagcgctc 1680
gtgagattgt ggagagctga ggaaccagtg ctgtgttgac cgagggttct tttgcaggag 1740
cgatgtacgt tgtcccttcc tctgggcttt gcggggggtgg gggtgggggt ggggggctgct 1800
gctgctgtgg ctgctcctcg tgaatcccttt tcttttcatc tttttgcatta cttttcttgc 1860
tttcatcacc aacctatttc tcatctgcct cctcccctac tgaaacccct gttttcctat 1920
tgccagagag agccatctgt cagaatatgg ctgtagttac tatcataatt caaatattcc 1980
ctttttaagta tttcttgacc caacaattac aaagaagcac tgctagatac aaacataaat 2040
aaaatgtgct aatgaggaat gcaaacagta ttaaaatgta aacttaaaat aggaaggaat 2100
ttgtggccaa agaaaaactg attacgattc tttcagtctt ttaaaatgca tcataaaact 2160
aaaaatgttt tgctacagaa taacagtttg gcttctttta tgatggcaaa gaagggtaat 2220
tacgattgtc atgaaaacac attgagcagc tgtatactgt taacaaaaga accattatat 2280
gagctcgtaa aagctggcac tttatatatt cacatttcct agaaattaga ttgtcttccc 2340
aattggagtc gataccaaaa gagactgttt tctattaatt gtggctttgt accagtcaat 2400
tttcagagcc aactggatca tctttctttt gggatatatc ccacatatta ttatgtggaa 2460
tttatatata ctgatttata tacactcaga agcatacagt ataatgttgc aaagagtaat 2520
tttttaagtc cccagagatc tgtaaattt taacgttaaa aaatgccagt ttcagtagat 2580
tctttcacca acgtacatct gtttagaaat tactgaaaga aatcataacg cataactctg 2640
cagtttattt tatgaaggct aattaatagg tacttactca ttaaagacta ggtctggtgc 2700
aaaatagaga aattggctgt tcgtatgttt gtacgatctc cagctcaagg caaatgatga 2760
tagacacatc caagaatact ggattagggt aatttggtcc tcaagaggca agttttaaa 2820
tcctgaagaa caaaacaatt aatcacagaa atacacttag catttaagta cattccagga 2880
agatgcttct taagtcaacc ccaaacagcc acctttcttt cactttctc ttttctgaca 2940
gactcaaatc aattcaacag tcatttacag agtgcccacc actggcccag tcctgttcca 3000
catttcactt taaattcatc ctaaattcta gatgagggta aactgccagg gcaaactaaa 3060
attgcaaagt tcagcaagtt tcagtgtgga tgacagccta aaacccttgg ttagattgag 3120
agaatttgta ttaacctgta aaacttgctt cgtcccaccc acaattaata acactaaaat 3180
aaagggtaat ttaataactg tcaaaccctg tgtatctgca ataatggtgg gagatttcat 3240
tgttcatgag                                                      3250
```

<210> 78
<211> 9212
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
     pCDGal4-DBD-MR sequence

<400> 78
```
agcactgcat aattctctta ctgtcatgcc atccgtaaga tgctttttctg tgactggtga 60
gtactcaacc aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc 120
gtcaatacgg gataataccg cgccacatag cagaactttta aaagtgctca tcattggaaa 180
acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta 240
acccactcgt gcacccaact gatcttcagc atctttttact ttcaccagcg tttctgggtg 300
```

```
agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg 360
aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat 420
gagcggatac atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt 480
tccccgaaaa gtgccacctg acgtcgacgg atcgggagat ctcccgatcc cctatggtgc 540
actctcagta caatctgctc tgatgccgca tagttaagcc agtatctgct ccctgcttgt 600
gtgttggagg tcgctgagta gtgcgcgagc aaaatttaag ctacaacaag gcaaggcttg 660
accgacaatt gcatgaagaa tctgcttagg gttaggcgtt ttgcgctgct tcgcgatgta 720
cgggccagat atacgcgttg acattgatta ttgactagtt attaatagta atcaattacg 780
gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaatggc 840
ccgcctggct gaccgcccaa cgacccccgc ccattgacgt caataatgac gtatgttccc 900
atagtaacgc caatagggac tttccattga cgtcaatggg tggagtattt acggtaaact 960
gcccacttgg cagtacatca agtgtatcat atgccaagta cgcccccctat tgacgtcaat 1020
gacggtaaat ggcccgcctg gcattatgcc cagtacatga ccttatggga ctttcctact 1080
tggcagtaca tctacgtatt agtcatcgct attaccatgg tgatgcggtt ttggcagtac 1140
atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca ccccattgac 1200
gtcaatggga gtttgttttg gcaccaaaat caacgggact ttccaaaatg tcgtaacaac 1260
tccgccccat tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta tataagcaga 1320
gctctctggc taactagaga acccactgct tactggctta tcgaaattaa tacgactcac 1380
tatagggaga cccaagctgg ctagcgttta aacttaagct tgccaccatg aagctactgt 1440
cttctatcga caagcatgc gatatttgcc gacttaaaaa gctcaagtgc tccaaagaaa 1500
aaccgaagtg cgccaagtgt ctgaagaaca actgggagtg tcgctactct cccaaaacca 1560
aaaggtctcc gctgactagg gcacatctga cagaagtgga atcaaggcta gaaagactgg 1620
aacagctatt tctactgatt tttcctcgag aagaccttga catgattttg aaaatggatt 1680
ctttacagga tataaaagca ttgttaacag gattatttgt acaagataat gtgaataaag 1740
atgccgtcac agatagattg gcttcagtgg agactgatat gcctctaaca ttgagacagc 1800
atagaataag tgcgacatca tcatcggaag agagtagtaa caaaggtcaa agacagttga 1860
ctgtatcgcc gggtaccttt gtggtgctta aaaaatgagc atgtggtgcc tcctatggtc 1920
caggttaata ccagactttg ataaagttct atggtagaac atgtttccct ggcacctgtc 1980
ctccctgcct gacagactca agctcttctt ttgttcttta ttttatttat ttatttttta 2040
agatggattc tcattctgtc gcccaggctg gagttcagtg acatgatatt ggctcactgc 2100
atcccaggtt tcaagcgatt ctcctgcctc aggctcacaa gtagctgggg ttacaggtgc 2160
ctgccaggac gcccaactaa attttgtatt tttagtagag acagggtttc atgatggtgg 2220
caaggctggt ctcaaattcc tgacctcaag tgatctgctt gcctcagcct cccaaagtct 2280
gggattacag gcatgggcca ctgcaccagg ccttcttttg tttttttatca cctcccccat 2340
ctccaggggc atcatgttat ttttctgcct gtggagtttc actgggcaag ctcctcttca 2400
aagaatcagt cagtggagct ttccccaact tttacagctg ctgtctgacc acaggtaatc 2460
taagaaactg ctttagatta atgaactggg acaatgtaat ggcaatttct aagtgatgtt 2520
tctgtagatt aagagggggcc ttattaatta aataccatga aacaaaaaaa aatgactaaa 2580
ataggaattc cagggacaat ttaaagggtg ctgggggcca ggtgattatt ttttagacag 2640
aagagtcatt ccaggtctct tacaacttac ttgctaggcc ttgccttact ggtctacac 2700
ttagattaac cctcgttatg agccgctaat aatgtgctca acttgtgatt tcatgaacag 2760
aatgtatttt accttattaa tatttttgga tgaaattttc cattttgatc aacataagaa 2820
gtgtatgtca tttttttttat ttttaatttc tttttttttt ttgagacaga gtctcactct 2880
gttgctcagg ctggagtata gtggtgggat ctcggctcac tgcagcctct gccgcccagg 2940
ttcaagcgat tctcctgcct cagcctcctg agtagctggg attacaggca cctgccacca 3000
agcaggacta attttttgtgt ttttagtaga aacggggttt caccatcttg gctaggctgg 3060
tcttgaactc ctgacctcgt gatccatctg cctcagtaag ccacctgcct tggcctccca 3120
aagtgctggg attacaggcg tgagccactg cgcccagccg tatctgtcat tttaacaata 3180
aaatccttaa ccctgcattc tcggaagaaa tgagcacatt aaccatattt ttcctataca 3240
cttgatcaat ttgttatcaa tttaaaaaat ctgatatcag gatttcatag aacgcaaact 3300
cctcctgcat ggttggagat ggcgaagtca gttgcccaga cttgttaagt tcaggatgca 3360
gcctgtgaaa ggagaggcaa tcctacctgg aagtaccttt gcccacttca cgacttggat 3420
catctgtttg cctgctaagc ggttgagcgt ggagagcaga ttttcggctg tatctggttt 3480
tgagctgtca tagcctgcat atacaatttc aggttcaatg ttttcaagga ccataacggg 3540
ggaaggtgtg agcgctcgtg agattgtgga gagctgagga accagtgctg tgttgaccga 3600
gggttctttt gcaggagcga tgtacgttgt cccttcctct gggctttgcg ggggtggggg 3660
tgggggtggg ggctgctgct gctgtggctg ctcctcgtga atccctttct tttcatcttt 3720
tgcattactt ttcttgcttt catcaccaac ctatttctca tctgcctcct cccctactga 3780
```

```
aacccctgtt ttcctattgc cagagagagc catctgtcag aatatggctg tagttactat 3840
cataattcaa atattccctt ttaagtatttt cttgacccaa caattacaaa gaagcactgc 3900
tagatacaaa cataaataaa atgtgctaat gaggaatgca aacagtatta aaatgtaaac 3960
ttaaaatagg aaggaatttg tggccaaaga aaaactgatt acgattcttt cagtctttta 4020
aaatgcatca taaaactaaa aatgttttgc tacagaataa cagtttggct tcttttatga 4080
tggcaaagaa gggtaattac gattgtcatg aaaacacatt gagcagctgt atactgttaa 4140
caaaagaacc attatatgag ctcgtaaaag ctggcacttt atatattcac atttcctaga 4200
aattagattg tcttcccaat tggagtcgat accaaaagag actgtttct attaattgtg 4260
gctttgtacc agtcaatttt cagagccaac tggatcatct ttcttttggg atatatccca 4320
catattatta tgtggaattt atatatactg atttatatac actcagaagc atacagtata 4380
atgttgcaaa gagtaatttt ttaagttccc agagatctgt aaattttaa cgttaaaaaa 4440
tgccagtttc agtagattct ttcaccaacg tacatctgtt tagaaattac tgaaagaaat 4500
cataacgcat aactctgcag tttattttat gaaggctaat taataggtac ttactcatta 4560
aagactaggt ctggtgcaaa atagagaaat tggctgttcg tatgtttgta cgatctccag 4620
ctcaaggcaa atgatgatag acacatccaa gaatactgga ttagggtaat ttggtcctca 4680
agaggcaagt ttttaaatcc tgaagaacaa aacaattaat cacagaaata cacttagcat 4740
ttaagtacat tccaggaaga tgcttcttaa gtcaacccca aacagccacc tttctttcac 4800
ttttctcttt tctgacagac tcaaatcaat tcaacagtca tttacagagt gcccaccact 4860
ggcccagtcc tgttccacat ttcacttttaa attcatccta aattctagat gagggtaaac 4920
tgccagggca aactaaaatt gcaaagttca gcaagtttca gtgtggatga cagcctaaaa 4980
cccttggtta gattgagaga atttgtatta acctgaaaa cttgcttcgt cccacccaca 5040
attaataaca ctaaaataaa gggtaattta ataactgtca aaccctgtgt atctgcaata 5100
atggtgggag atttcattgt tcatgaggcg gccgctcgag tctagagggc ccgtttaaac 5160
ccgctgatca gcctcgactg tgccttctag ttgccagcca tctgttgttt gcccctcccc 5220
cgtgccttcc ttgaccctgg aaggtgccac tcccactgtc ctttcctaat aaaatgagga 5280
aattgcatcg cattgtctga gtaggtgtca ttctattctg ggggtgggg tggggcagga 5340
cagcaagggg gaggattggg aagacaatag caggcatgct ggggatgcgg tgggctctat 5400
ggcttctgag gcggaaagaa ccagctgggg ctctaggggg tatccccacg cgccctgtag 5460
cggcgcatta gcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag 5520
cgccctagcg cccgctcctt tcgctttctt cccttccttt ctcgccacgt tcgccggctt 5580
tccccgtcaa gctctaaatc ggggggctccc tttagggttc cgatttagtg ctttacggca 5640
cctcgacccc aaaaaacttg attagggtga tggttcacgt agtgggccat cgccctgata 5700
gacggttttt cgccctttga cgttggagtc cacgttcttt aatagtggac tcttgttcca 5760
aactggaaca acactcaacc ctatctcggt ctattctttt gatttataag ggattttgcc 5820
gatttcggcc tattggttaa aaaatgagct gatttaacaa aaatttaacg cgaattaatt 5880
ctgtggaatg tgtgtcagtt agggtgtgga aagtccccag gctccccagc aggcagaagt 5940
atgcaaagca tgcatctcaa ttagtcagca accaggtgtg gaaagtcccc aggctcccca 6000
gcaggcagaa gtatgcaaag catgcatctc aattagtcag caaccatagt cccgccccta 6060
actccgccca tcccgcccct aactccgccc agttccgccc attctccgcc ccatggctga 6120
ctaattttt ttatttatgc agaggccgag gccgcctctg cctctgagct attccagaag 6180
tagtgaggag gcttttttgg aggcctaggc ttttgcaaaa agctcccggg agcttgtata 6240
tccatttcg gatctgatca gcacgtgatg aaaaagcctg aactcaccgc gacgtctgtc 6300
gagaagtttc tgatcgaaaa gttcgacagc gtctccgacc tgatgcagct ctcggagggc 6360
gaagaatctc gtgctttcag cttcgatgta ggagggcgtg gatatgtcct gcgggtaaat 6420
agctgcgccg atggtttcta caaagatcgt tatgtttatc ggcactttgc atcggccgcg 6480
ctcccgattc cggaagtgct tgacattggg gaattcagcg agagcctgac ctattgcatc 6540
tcccgccgtg cacagggtgt cacgttgcaa gacctgcctg aaaccgaact gcccgctgtt 6600
ctgcagccgg tcgcggaggc catggatgcg atcgctgcgg ccgatcttag ccagacgagc 6660
gggttcggcc cattcggacc gcaaggaatc ggtcaataca ctacatggcg tgatttcata 6720
tgcgcgattg ctgatcccca tgtgtatcac tggcaaactg tgatggacga caccgtcagt 6780
gcgtccgtcg cgcaggctct cgatgagctg atgctttggg ccgaggactg ccccgaagtc 6840
cggcacctcg tgcacgcgga tttcggctcc aacaatgtcc tgacggacaa tggccgcata 6900
acagcggtca ttgactggag cgaggcgatg ttcggggatt cccaatacga ggtcgccaac 6960
atcttcttct ggaggccgtg gttggcttgt atggagcagc agacgcgcta cttcgagcgg 7020
aggcatccgg agcttgcagg atcgccgcgg ctccgggcgt atatgctccg cattggtctt 7080
gaccaactct atcagagctt ggttgacggc aatttcgatg atgcagcttg ggcgcagggt 7140
cgatgcgacg caatcgtccg atccggagcc gggactgtcg ggcgtacaca aatcgcccgc 7200
agaagcgcgg ccgtctggac cgatggctgt gtagaagtac tcgccgatag tggaaaccga 7260
```

```
cgccccagca ctcgtccgag ggcaaaggaa tagcacgtgc tacgagattt cgattccacc 7320
gccgccttct atgaaaggtt gggcttcgga atcgttttcc gggacgccgg ctggatgatc 7380
ctccagcgcg gggatctcat gctggagttc ttcgcccacc ccaacttgtt tattgcagct 7440
tataatggtt acaaataaag caatagcatc acaaatttca caaataaagc attttttttca 7500
ctgcattcta gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctgtataccg 7560
tcgacctcta gctagagctt ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt 7620
tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa agcctggggt 7680
gcctaatgag tgagctaact cacattaatt gcgttgcgct cactgcccgc tttccagtcg 7740
ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag aggcggtttg 7800
cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg 7860
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga atcaggggat 7920
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 7980
gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacaa aaatcgacgc 8040
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga 8100
agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt 8160
ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg 8220
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc 8280
gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg 8340
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc 8400
ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat ctgcgctctg 8460
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc 8520
gctggtagcg gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa 8580
gaagatcctt tgatctttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa 8640
gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa 8700
tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag ttaccaatgc 8760
ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat agttgcctga 8820
ctccccgtcg tgtagataac tacgatacgg gagggcttac catctggccc cagtgctgca 8880
atgataccgc gagacccacg ctcaccggct ccagatttat cagcaataaa ccagccagcc 8940
ggaagggccg agcgcagaag tggtcctgca actttatccg cctccatcca gtctattaat 9000
tgttgccggg aagctagagt aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc 9060
attgctacag gcatcgtggt gtcacgctcg tcgtttggta tggcttcatt cagctccggt 9120
tcccaacgat caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc 9180
ttcggtcctc cgatcgttgt cagaagtaag tt                                  9212
```

<210> 79
<211> 10517
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
     pKI-Gal-DBD-MR sequence

<220>
<221> modified_base
<222> (9953)..(9953)
<223> any nucleotide

<400> 79
```
atcggacgat tgcgtcgcat cgaccctgcg cccaagctgc atcatcgaaa ttgccgtcaa 60
ccaagctctg atagagttgg tcaagaccaa tgcggagcat atacgcccgg agccgcggcg 120
atcctgcaag ctccggatgc ctccgctcga agtagcgcgt ctgctgctcc atacaagcca 180
accacggcct ccagaagaag atgttggcga cctcgtattg ggaatccccg aacatcgcct 240
cgctccagtc aatgaccgct gttatgcggc cattgtccgt caggacattg ttggagccga 300
aatccgcgtg cacgaggtgc cggacttcgg ggcagtcctc ggcccaaagc atcagctcat 360
cgagagcctg cgcgacggac gcactgacgg tgtcgtccat cacagtttgc cagtgataca 420
catggggatc agcaatcgcg catatgaaat cacgccatgt agtgtattga ccgattcctt 480
```

```
gcggtccgaa tgggccgaac ccgctcgtct ggctaagatc ggccgcagcg atcgcatcca 540
tggcctccgc gaccggctgc agaacagcgg gcagttcggt ttcaggcagg tcttgcaacg 600
tgacaccctg tgcacggcgg gagatgcaat aggtcaggct ctcgctgaat ccccaatgt 660
caagcacttc cggaatcggg agcgcggccg atgcaaagtg ccgataaaca taacgatctt 720
tgtagaaacc atcggcgcag ctatttaccc gcaggacata tccacgccct cctacatcga 780
agctgaaagc acgagattct tcgccctccg agagctgcat caggtcggag acgctgtcga 840
actttcgat cagaaacttc tcgacagacg tcgcggtgag ttcaggcttt ttcatgcagg 900
tcgaaggccc ggagatgagg aagaggagaa cagcgcggca gacgtgcgct tttgaagcgt 960
gcagaatgcc gggtccggag gaccttcgcg cccgccccgc ccctgagccc gccctgagc 1020
ccgcccccgg acccaccect tcccagcctc tgagcccaga aagcgaagga gccaagctgc 1080
tattggccgc tgccccaaag gcctacccgc ttccattgct cagcggtgct tgtccatctg 1140
cacgagacta gtgagacgtg ctacttccat ttgtcacgtc ctgcacgacg cgagctgcgg 1200
ggcgggggg aacttcctga ctaggggagg agtagaaggt ggcgcgaagg ggccaccaaa 1260
gaagggagcc ggttggcgct accggtggat gtggaatgtg tgcgaggcca gaggccactt 1320
gtgatgcgcc aagtgccagc ggggctgcta aagcgcatgc tccagactgc cttgggaaaa 1380
gcgcctcccc tacccggtag aattcctgca gcccggggtt gacattgatt attgactagt 1440
tattaatagt aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt 1500
acataactta cggtaaatgg cccgcctggc tgaccgccca acgacccccg cccattgacg 1560
tcaataatga cgtatgttcc catagtaacg ccaatagggA ctttccattg acgtcaatgg 1620
gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt 1680
acgccccta ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc ccagtacatg 1740
accttatggg actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg 1800
gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt 1860
ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac 1920
tttccaaaat gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag gcgtgtacgg 1980
tgggaggtct atataagcag agctctctgg ctaactagag aacccactgc ttactggctt 2040
atccggatcc actagttcta gcgtttaaac ttaagcttgc caccatgaag ctactgtctt 2100
ctatcgaaca agcatgcgat atttgccgac ttaaaaagct caagtgctcc aaagaaaaac 2160
cgaagtgcgc caagtgtctg aagaacaact gggagtgtcg ctactctccc aaaaccaaaa 2220
ggtctccgct gactagggca catctgacag aagtggaatc aaggctagaa agactggaac 2280
agctatttct actgattttt cctcgagaag accttgacat gattttgaaa atggattctt 2340
tacaggatat aaaagcattg ttaacaggat tatttgtaca agataatgtg aataaagatg 2400
ccgtcacaga tagattggct tcagtggaga ctgatatgcc tctaacattg agacagcata 2460
gaataagtgc gacatcatca tcggaagaga gtagtaacaa aggtcaaaga cagttgactg 2520
tatcgccggg tacctttgtg gtgcttaaaa aatgagcatg tggtgcctcc tatggtccag 2580
gttaatacca gactttgata aagttctatg gtagaacatg tttccctggc acctgtcctc 2640
cctgcctgac agactcaagc tcttctttg ttctttattt tatttatta tttttaaga 2700
tggattctca ttctgtcgcc caggctggag ttcagtgaca tgatattggc tcactgcatc 2760
ccaggtttca agcgattctc ctgcctcagg ctcacaagta gctggggtta caggtgcctg 2820
ccaggacgcc caactaaatt ttgtattttt agtagagaca gggtttcatg atggtggcaa 2880
ggctggtctc aaattcctga cctcaagtga tctgcttgcc tcagcctccc aaagtctggg 2940
attacaggca tgggccactg caccaggcct tctttgttt tttatcacct cccccatctc 3000
cagggggcatc atgttatttt tctgcctgtg gagtttcact gggcaagctc ctcttcaaag 3060
aatcagtcag tggagctttc cccaacttttt acagctgctg tctgaccaca ggtaatctaa 3120
gaaactgctt tagattaatg aactgggaca atgtaatggc aatttctaag tgatgtttct 3180
gtagattaag aggggcctta ttaattaaat accatgaaac aaaaaaaaat gactaaaata 3240
ggaattccag ggacaattta aagggtgctg ggggccaggt gattattttt tagacagaag 3300
agtcattcca ggtctcttac aacttacttg ctaggccttg ccttacttgg tctacactta 3360
gattaaccct cgttatgagc cgctaataat gtgctcaact tgtgatttca tgaacagaat 3420
gtattttacc ttattaatat ttttggatga aattttccat tttgatcaac ataagaagtg 3480
tatgtcattt tttttatttt taatttcttt ttttttttg agacagagtc tcactctgtt 3540
gctcaggctg gagtatagtg gtgggatctc ggctcactgc agcctctgcc gcccaggttc 3600
aagcgattct cctgcctcag cctcctgagt agctgggatt acaggcacct gccaccaagc 3660
aggactaatt tttgtgtttt tagtagaaac ggggtttcac catcttggct aggctggtct 3720
tgaactcctg acctcgtgat ccatctgcct cagtaagcca cctgccttgg cctcccaaag 3780
tgctgggatt acaggcgtga gccactgcgc ccagccgtat ctgtcatttt aacaataaaa 3840
tccttaaccc tgcattctcg gaagaaatga gcacattaac catatttttc ctatacactt 3900
gatcaatttg ttatcaattt aaaaaatctg atatcaggat ttcatagaac gcaaactcct 3960
```

```
cctgcatggt tggagatggc gaagtcagtt gcccagactt gttaagttca ggatgcagcc 4020
tgtgaaagga gaggcaatcc tacctggaag tacctttgcc cacttcacga cttggatcat 4080
ctgtttgcct gctaagcggt tgagcgtgga gagcagattt tcggctgtat ctggttttga 4140
gctgtcatag cctgcatata caatttcagg ttcaatgttt tcaaggacca taacggggga 4200
aggtgtgagc gctcgtgaga ttgtggagag ctgaggaacc agtgctgtgt tgaccgaggg 4260
ttcttttgca ggagcgatgt acgttgtccc ttcctctggg ctttgcgggg gtgggggtgg 4320
gggtgggggc tgctgctgct gtggctgctc ctcgtgaatc cctttctttt catcttttgc 4380
attacttttc ttgctttcat caccaaccta tttctcatct gcctcctccc ctactgaaac 4440
ccctgttttc ctattgccag agagagccat ctgtcagaat atggctgtag ttactatcat 4500
aattcaaata ttcccttta agtatttctt gacccaacaa ttacaaagaa gcactgctag 4560
atacaaacat aaataaaatg tgctaatgag gaatgcaaac agtattaaaa tgtaaactta 4620
aaataggaag gaatttgtgg ccaaagaaaa actgattacg attctttcag tctttaaaa 4680
tgcatcataa aactaaaaat gttttgctac agaataacag tttggcttct tttatgatgg 4740
caaagaaggg taattacgat tgtcatgaaa acacattgag cagctgtata ctgttaacaa 4800
aagaaccatt atatgagctc gtaaaagctg gcactttata tattcacatt tcctagaaat 4860
tagattgtct tcccaattgg agtcgatacc aaaagagact gttttctatt aattgtggct 4920
ttgtaccagt caattttcag agccaactgg atcatctttc ttttgggata tatcccacat 4980
attattatgt ggaatttata tatactgatt tatatacact cagaagcata cagtataatg 5040
ttgcaaagag taattttta agttcccaga gatctgtaaa tttttaacgt taaaaaatgc 5100
cagtttcagt agattctttc accaacgtac atctgtttag aaattactga aagaaatcat 5160
aacgcataac tctgcagttt attttatgaa ggctaattaa taggtactta ctcattaaag 5220
actaggtctg gtgcaaaata gagaaattgg ctgttcgtat gtttgtacga tctccagctc 5280
aaggcaaatg atgatagaca catccaagaa tactggatta gggtaatttg gtcctcaaga 5340
ggcaagtttt taaatcctga agaacaaac aattaatcac agaaatacac ttagcattta 5400
agtacattcc aggaagatgc ttcttaagtc aaccccaaac agccaccttt ctttcacttt 5460
tctctttct gacagactca aatcaattca acagtcattt acagagtgcc caccactggc 5520
ccagtcctgt tccacatttc actttaaatt catcctaaat tctagatgag ggtaaactgc 5580
cagggcaaac taaaattgca aagttcagca agtttcagtg tggatgacag cctaaaaccc 5640
ttggttagat tgagagaatt tgtattaacc tgtaaaactt gcttcgtccc acccacaatt 5700
aataacacta aaataaaggg taatttaata actgtcaaac cctgtgtatc tgcaataatg 5760
gtgggagatt tcattgttca tgaggcggcc gcttgatctg atctgagtca ctgaatggaa 5820
tactgtccgc aagggcgggg gctattcttg tccatctggg cctacggaac cagcacccat 5880
cgccaaactc ttcatcctct tcctcaatct cgctttctct ctcgctcgct tttttttctct 5940
tctcctttt tttttttttt tttttttttc ccccgcaaaa ggaggggaga ggggtaaaa 6000
aaatgctgca ctgtgcggca aggccggtga ataaacgact cgaatccaat caccccccgc 6060
cgttcctaaa gttgcctttt atggctcgag tggccgctgt ggcttcctat aaaacccggc 6120
ggcgcaaccc ccagccactg tcaaatcaag ctaacgccac catggtgagc aagggcgagg 6180
agctgttcac cggggtggtg cccatcctgg tcgagctgga cggcgacgta aacggccaca 6240
agttcagcgt gtccggcgag ggcgagggcg atgccaccta cggcaagctg accctgaagt 6300
tcatctgcac caccggcaag ctgcccgtgc cctggcccac cctcgtgacc accttcggct 6360
acggcctgca gtgcttcgcc cgctaccccg accacatgaa gcagcacgac ttcttcaagt 6420
ccgccatgcc cgaaggctac gtccaggagc gcaccatctt cttcaaggac gacggcaact 6480
acaagacccg cgccgaggtg aagttcgagg gcgacaccct ggtgaaccgc atcgagctga 6540
agggcatcga cttcaaggag gacggcaaca tcctggggca caagctggag tacaactaca 6600
acagccacaa cgtctatatc atggccgaca gcagaagaa cggcatcaag gtgaacttca 6660
agatccgcca caacatcgag gacggcagcg tgcagctcgc cgaccactac cagcagaaca 6720
cccccatcgg cgacggcccc gtgctgctgc ccgacaacca ctacctgagc taccagtccg 6780
ccctgagcaa agaccccaac gagaagcgcg atcacatggt cctgctggag ttcgtgaccg 6840
ccgccgggat cactctcggc atggacgagc tgtacaagaa cttgtttatt gcagcttata 6900
atggttacaa ataaagcaat agcatcacaa atttcacaaa taaagcattt ttttcactgc 6960
attctagttg tggtttgtcc aaactcatca atgtatctta tcatgtcgtg gcctgcatgg 7020
cccggccgca ccgcggtgga gctccaattc gccctatagt gagtcgtatt acgcgcgctc 7080
actggccgtc gttttacaac gtcgtgactg ggaaaaccct ggcgttaccc aacttaatcg 7140
ccttgcagca catccccctt tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg 7200
cccttcccaa cagttgcgca gcctgaatgg cgaatgggac gcgccctgta gcggcgcatt 7260
aagcgcggcg ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc 7320
gcccgctcct ttcgctttct tcccttcctt tctcgccacg ttcgccggct ttccccgtca 7380
agctctaaat cggggggctcc ctttagggtt ccgatttagt gctttacggc acctcgaccc 7440
```

```
caaaaaactt gattagggtg atggttcacg tagtgggcca tcgccctgat agacggtttt 7500
tcgccctttg acgttggagt ccacgttctt taatagtgga ctcttgttcc aaactggaac 7560
aacactcaac cctatctcgg tctattcttt tgatttataa gggatttgc cgatttcggc 7620
ctattggtta aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt 7680
aacgcttaca atttaggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta 7740
tttttctaaa tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt 7800
caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc 7860
tttttttgcgg cattttgcct tcctgttttt gctcacccag aaacgctggt gaaagtaaaa 7920
gatgctgaag atcagttggg tgcacgagtg ggttacatcg aactggatct caacagcggt 7980
aagatccttg agagttttcg ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt 8040
ctgctatgtg gcgcggtatt atcccgtatt gacgccgggc aagagcaact cggtcgccgc 8100
atacactatt ctcagaatga cttggttgag tactcaccag tcacagaaaa gcatcttacg 8160
gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg 8220
gccaacttac ttctgacaac gatcggagga ccgaaggagc taaccgcttt tttgcacaac 8280
atggggggatc atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca 8340
aacgacgagc gtgacaccac gatgcctgta gcaatggcaa caacgttgcg caaactatta 8400
actggcgaac tacttactct agcttcccgg caacaattaa tagactggat ggaggcggat 8460
aaagttgcag gaccacttct gcgctcggcc cttccggctg gctggtttat tgctgataaa 8520
tctggagccg gtgagcgtgg gtctcgcggt atcattgcag cactggggcc agatggtaag 8580
ccctcccgta tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat 8640
agacagatcg ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt 8700
tactcatata tactttagat tgatttaaaa cttcatttt aatttaaaag gatctaggtg 8760
aagatccttt ttgataatct catgaccaaa atcccttaac gtgagttttc gttccactga 8820
gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atccttttt tctgcgcgta 8880
atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa 8940
gagctaccaa ctcttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact 9000
gtccttctag tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca 9060
tacctcgctc tgctaatcct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt 9120
accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg 9180
ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag 9240
cgtgagctat gagaaagcgc cacgcttccc gaagggagaa aggcggacag gtatccggta 9300
agcggcaggg tcggaacagg agagcgcacg agggagcttc caggggggaaa cgcctggtat 9360
ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg 9420
tcagggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc 9480
ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc tgtggataac 9540
cgtattaccg cctttgagtg agctgatacc gctcgccgca gccgaacgac cgagcgcagc 9600
gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct ccccgcgcgt 9660
tggccgattc attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag 9720
cgcaacgcaa ttaatgtgag ttagctcact cattaggcac cccaggcttt acactttatg 9780
cttccggctc gtatgttgtg tggaattgtg agcggataac aatttcacac aggaaacagc 9840
tatgaccatg attacgccaa gcgcgcaatt aaccctcact aaagggaaca aaagctgggt 9900
accgggcccc ccctcgaggt cgacggtatc gataagcttc cagtggatca ganaaacaat 9960
gatcccacga gagatctata gatctataga tcatgagtgg aggaatgagc tggcccttaa 10020
tttggttttg cttgtttaaa ttatgatatc caactatgaa acattatcat aaagcaatag 10080
taaagagcct tcagtaaaga gcaggcattt atctaatccc accccaccc ccacccccgt 10140
agctccaatc cttccattca aaatgtaggt actctgttct caccttctt aacaaagtat 10200
gacagtaaaa aacttccatt ttagtggaca tctttattgt ttaatagatc atcaatttat 10260
gcagacttac agctgggatt tcatcagatt gccatgctga gtcaagaaca gtgagtgacg 10320
aagctaacca gaggctacat acgtcagaga gagagctcag cctttacagc tcacttcctt 10380
tctcaggcat tcctttgccc tcggacgagt gctggggcgt cggttccac tatcggcgag 10440
tacttctaca cagccatcgg tccagacggc cgcgcttctg cgggcgattt gtgtacgccc 10500
gacagtcccg gctccgg                                              10517
```

<210> 80
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 80
gatctagaga gccacaccc atgttgtctc ac   32

<210> 81
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 81
gagcggccgc ttgccaaggg ccctgtgaag cagg   34

<210> 82
<211> 5351
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      pKI-CMV-SD sequence

<400> 82
tcccttagt gagggttaat tgcgcgcttg gcgtaatcat ggtcatagct gtttcctgtg 60
tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat aaagtgtaaa 120
gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgctc actgcccgct 180
ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgcggggaga 240
ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct gcgctcggtc 300
gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt atccacagaa 360
tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc aggaaccgt 420
aaaaaggccg cgttgctggc gtttttccat aggctccgcc ccctgacga gcatcacaaa 480
aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt 540
cccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg 600
tccgccttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc 660
agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaacccc cgttcagccc 720
gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta 780
tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct 840
acagagttct tgaagtggtg gcctaactac ggctacacta gaaggacagt atttggtatc 900
tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa 960
caaaccaccg ctggtagcgg tggtttttt gtttgcaagc agcagattac gcgcagaaaa 1020
aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa 1080
aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt 1140
ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac 1200
agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc 1260
atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt accatctggc 1320

```
cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata 1380
aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc 1440
cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc 1500
aacgttgttg ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca 1560
ttcagctccg gttcccaacg atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa 1620
gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca 1680
ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt 1740
tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt 1800
tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg 1860
ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga 1920
tccagttcga tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc 1980
agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg 2040
acacggaaat gttgaatact catactcttc ctttttcaat attattgaag catttatcag 2100
ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg 2160
gttccgcgca catttccccg aaaagtgcca cctaaattgt aagcgttaat attttgttaa 2220
aattcgcgtt aaattttttgt taaatcagct cattttttaa ccaataggcc gaaatcggca 2280
aaatccctta taaatcaaaa gaatagaccg agatagggtt gagtgttgtt ccagtttgga 2340
acaagagtcc actattaaag aacgtggact ccaacgtcaa agggcgaaaa accgtctatc 2400
agggcgatgg cccactacgt gaaccatcac cctaatcaag ttttttgggg tcgaggtgcc 2460
gtaaagcact aaatcggaac cctaaaggga gcccccgatt tagagcttga cggggaaagc 2520
cggcgaacgt ggcgagaaag gaagggaaga aagcgaaagg agcgggcgct agggcgctgg 2580
caagtgtagc ggtcacgctg cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac 2640
agggcgcgtc ccattcgcca ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg 2700
cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga ttaagttggg 2760
taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagtgag cgcgcgtaat 2820
acgactcact atagggcgaa ttggagctcc accgcggtgg cggccgctct agaactagtg 2880
gatcccccgg gaccggtccc cccactcacc tgccagtaag cagtgggttc tctagttagc 2940
cagagagctc tgcttatata gacctcccac cgtacacgcc taccgcccat ttgcgtcaat 3000
ggggcggagt tgttacgaca ttttggaaag tcccgttgat tttggtgcca aaacaaactc 3060
ccattgacgt caatgggtg gagacttgga aatccccgtg agtcaaaccg ctatccacgc 3120
ccattgatgt actgccaaaa ccgcatcacc atggtaatag cgatgactaa tacgtagatg 3180
tactgccaag taggaaagtc ccataaggtc atgtactggg cataatgcca ggcgggccat 3240
ttaccgtcat tgacgtcaat aggggcgta cttggcatat gatacacttg atgtactgcc 3300
aagtgggcag tttaccgtaa atagtccacc cattgacgtc aatggaaagt ccctattggc 3360
gttactatgg gaacatacgt cattattgac gtcaatgggc gggggtcgtt gggcggtcag 3420
ccaggcgggc catttaccgt aagttatgta acgcggaact ccatatatgg gctatgaact 3480
aatgaccccg taattgatta ctattaataa ctagtcaata atcaatgtca acgcccgggc 3540
tgcaggaatt ctaccgggta ggggaggcgc ttttcccaag gcagtctgga gcatgcgctt 3600
tagcagcccc gctggcactt ggcgcatcac aagtggcctc tggcctcgca cacattccac 3660
atccaccggt agcgccaacc ggctcccttc tttggtggcc ccttcgcgcc accttctact 3720
cctccctag tcaggaagtt cccccccgcc ccgcagctcg cgtcgtgcag gacgtgacaa 3780
atggaagtag cacgtctcac tagtctcgtg cagatggaca agcaccgctg agcaatggaa 3840
gcgggtaggc ctttggggca gcggccaata gcagcttggc tccttcgctt tctgggctca 3900
gaggctggga aggggtgggt ccggggcgg gctcaggggc gggctcaggg gcggggcggg 3960
cgcgaaggtc ctccggaccc ggcattctgc acgcttcaaa agcgcacgtc tgccgcgctg 4020
ttctcctctt cctcatctcc gggccttcga cctgcatgaa aaagcctgaa ctcaccgcga 4080
cgtctgtcga gaagtttctg atcgaaaagt tcgacagcgt ctccgacctg atgcagctct 4140
cggagggcga agaatctcgt gctttcagct tcgatgtagg agggcgtgga tatgtcctgc 4200
gggtaaatag ctgcgccgat ggtttctaca aagatcgtta tgtttatcgg cactttgcat 4260
cggccgcgct cccgattccg gaagtgcttg acattgggga attcagcgag agcctgacct 4320
attgcatctc ccgccgtgca cagggtgtca cgttgcaaga cctgcctgaa accgaactgc 4380
ccgctgttct gcagccggtc gcggaggcca tggatgcgat cgctgcggcc gatcttagcc 4440
agacgagcgg gttcggccca ttcggaccgc aaggaatcgg tcaatacact acatggcgtg 4500
atttcatatg cgcgattgct gatccccatg tgtatcactg gcaaactgtg atggacgaca 4560
ccgtcagtgc gtccgtcgcg caggctctcg atgagctgat gctttgggcc gaggactgcc 4620
ccgaagtccg gcacctcgtg cacgcggatt cggctccaa caatgtcctg acggacaatg 4680
gccgcataac agcggtcatt gactggagcg aggcgatgtt cggggattcc caatacgagg 4740
tcgccaacat cttcttctgg aggccgtggt tggcttgtat ggagcagcag acgcgctact 4800
```

```
tcgagcggag gcatccggag cttgcaggat cgccgcggct ccgggcgtat atgctccgca 4860
ttggtcttga ccaactctat cagagcttgg ttgacggcaa tttcgatgat gcagcttggg 4920
cgcagggtcg atgcgacgca atcgtccgat ccggagccgg gactgtcggg cgtacacaaa 4980
tcgcccgcag aagcgcggcc gtctggaccg atggctgtgt agaagtactc gccgatagtg 5040
gaaaccgacg ccccagcact cgtccgaggg caaaggaatg agctcgctga tcagcctcga 5100
ctgtgccttc tagttgccag ccatctgttg tttgcccctc ccccgtgcct tcccattgtc 5160
tgagtaggtg tcattctatt ctggggggtg gggtggggca ggacagcaag ggggaggatt 5220
gggaagacaa tagcaggcat gctggggatg cggtgggctc tatggcttct gaggcggaaa 5280
gaaccagctg ggggatatca agcttatcga taccgtcgac ctcgagggg ggcccggtac 5340
ccagctttg t                                                     5351


<210> 83
<211> 9131
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      pKI-CMV-SD-Vanilloid sequence

<400> 83
cgatcgctgc ggccgatctt agccagacga gcgggttcgg cccattcgga ccgcaaggaa 60
tcggtcaata cactacatgg cgtgatttca tatgcgcgat tgctgatccc catgtgtatc 120
actggcaaac tgtgatggac gacaccgtca gtgcgtccgt cgcgcaggct ctcgatgagc 180
tgatgctttg ggccgaggac tgccccgaag tccggcacct cgtgcacgcg gatttcggct 240
ccaacaatgt cctgacggac aatggccgca taacagcggt cattgactgg agcgaggcga 300
tgttcgggga ttcccaatac gaggtcgcca acatcttctt ctggaggccg tggttggctt 360
gtatggagca gcagacgcgc tacttcgagc ggaggcatcc ggagcttgca ggatcgccgc 420
ggctccgggc gtatatgctc cgcattggtc ttgaccaact ctatcagagc ttggttgacg 480
gcaatttcga tgatgcagct tgggcgcagg gtcgatgcgc agcgaatcgtc cgatccggag 540
ccgggactgt cgggcgtaca caaatcgccc gcagaagcgc ggccgtctgg accgatggct 600
gtgtagaagt actcgccgat agtggaaacc gacgccccag cactcgtccg agggcaaagg 660
aatgagctcg ctgatcagcc tcgactgtgc cttctagttg ccagccatct gttgtttgcc 720
cctccccccgt gccttcccat tgtctgagta ggtgtcattc tattctgggg ggtggggtgg 780
ggcaggacag caaggggggag gattgggaag acaatagcag gcatgctggg gatgcggtgg 840
gctctatggc ttctgaggcg gaaagaacca gctgggggat atcaagctta tcgataccgt 900
cgacctcgag ggggggcccg gtaccagct tttgttccct ttagtgaggg ttaattgcgc 960
gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg ctcacaattc 1020
cacacaacat acgagccgga agcataaagt gtaaagcctg gggtgcctaa tgagtgagct 1080
aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc 1140
agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt gggcgctctt 1200
ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg ctgcggcga gcggtatcag 1260
ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 1320
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt 1380
tccataggct ccgcccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc 1440
gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc ctcgtgcgct 1500
ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg 1560
tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca 1620
agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta ccggtaact 1680
atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta 1740
acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtggccta 1800
actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag ccagttacct 1860
tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt agcggtggtt 1920
tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga 1980
tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttggtca 2040
tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat 2100
caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg 2160
```

```
cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc cccgtcgtgt 2220
agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 2280
acccacgctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc 2340
gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag 2400
ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 2460
tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa 2520
ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc ggtcctccga 2580
tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcactgcata 2640
attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca 2700
agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 2760
ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg 2820
ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg 2880
cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag 2940
gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatac 3000
tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggataca 3060
tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 3120
tgccacctaa attgtaagcg ttaatatttt gttaaaattc gcgttaaatt tttgttaaat 3180
cagctcattt tttaaccaat aggccgaaat cggcaaaatc ccttataaat caaaagaata 3240
gaccgagata gggttgagtg ttgttccagt ttggaacaag agtccactat taaagaacgt 3300
ggactccaac gtcaaagggc gaaaaaccgt ctatcagggc gatggcccac tacgtgaacc 3360
atcaccctaa tcaagttttt tggggtcgag gtgccgtaaa gcactaaatc ggaaccctaa 3420
agggagcccc cgatttagag cttgacgggg aaagccggcg aacgtggcga aaaggaagg 3480
gaagaaagcg aaaggagcgg cgctagggc gctggcaagt gtagcggtca cgctgcgcgt 3540
aaccaccaca cccgccgcgc ttaatgcgcc gctacagggc gcgtcccatt cgccattcag 3600
gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct tcgctattac gccagctggc 3660
gaaagggga tgtgctgcaa ggcgattaag ttgggtaacg ccagggtttt cccagtcacg 3720
acgttgtaaa acgacggcca gtgagcgcgc gtaatacgac tcactatagg gcgaattgga 3780
gctccaccgc ggtggcggcc gcttgccaag ggcctgtga agcagggtac cccaaaactc 3840
caacttgctg ctgcccccag cccctagaa ctgaagtgtt tccaggactg aacctctcac 3900
cccagaggcc tcagaggtga gcaggcctag cagggcccag gcagtcctgc ggccactgac 3960
gtcggcctgg gagggtggac aggcccccgg ctcagagcca tctgtggttc tgggttagac 4020
ccatccctcc tcgtctctgg gcctcaggtg cctcctctgc agggtgaggt tgataaccac 4080
cgcgctgcct ccctccccac acagaacaag gaaaacggcc ggggtggctg tcgtgggctg 4140
aattcaggga ggatctcggt gatgaggatg gctcagtcca gaagagccaa gcagactgtg 4200
tccctggaat agggtgagcg gcagagactg gcccacccac acccacaccg cactgcgggg 4260
ccctgctgtg caacgggtcc acacatcatc agccaggtgc ccggcctgcc tttgctcttc 4320
aacctcgctg ccacagccac tgagcttata cgagattttg gattttgtgc tttttgtcat 4380
ttttatgaat caacatctgt caagcatcct ggccttgaat gtttaggcgt gtgtgtgtgt 4440
gtgtgtgtgt gtgtgtctca ctctgtcgcc cagactagag tgcaatgggg caatctctgt 4500
tcactgcaac ctccacctcc caggttcaag cgattctcct ccctcagcct cccaagtagc 4560
tgggattaca ggcgcccgcc accatgcgtg gctaatgttt gtattttag tagagacagg 4620
gaacgtttag gctttaagtg atttggagct ggttggtctt ctaggcctaa acacctagag 4680
atcagctggc cgctggcagc agcatgtgta gttggacaca tgagctgatg ctattttcca 4740
ctcccatcct gccccatccc catgcttccc ctccctctgg ctgggcatcc ccactccctg 4800
gcacccaggg gtggcccata catccctccc atagcctcta ttcccatcgg gtctccacct 4860
aagggctttt tcctctgtct acccttgcac aagtcctatc accactcaga tgctacttcc 4920
cctcaggatc cagcactgat gccgcctcct ccatgaagcc cccagatccc tgccctgctc 4980
atgctccccg ctgtcctcca cacagtggcc cagcccactg accgcctgtc ccacacagcc 5040
taccatgtgg tccagcaatt gctatctgtc actgatttat ctctcccgca ccctgtgacc 5100
tcctcctggg cagggacctt tcctggtcca tctccttatg aagaagctgt actggcctgg 5160
gcgcgtggct cacgcctgta atcccagcac tttgggaggc caaggcaggt ggatcacctg 5220
aggtcaggag ttcgagacca gcctggccaa catggtgaaa tcccatctct actaaaaaaa 5280
atacaaaaat tagccgggcg tggtgtcggg cgcctgtaat cccagctact cgggaggctg 5340
aggcaggaga atcgcttgaa cccaggatgt agaggttaca gtgagctgag actgtgccac 5400
tgcactccag cctgggcaac agagtgaaac tctgtctcaa aaaaaaaaa aaaaaaaaa 5460
aaagaagcag cagcagcagc agcagcagca gctgcacaca tggggaaata agagcaatca 5520
gaccccttcta attgccattt ccataggaag ttagggagat tactgttcca gcagtggcta 5580
cattccggaa aacagccgca gaagaaatga gaggcttctc aagtcctgca ttttctttt 5640
```

```
tttttttttt gaaatatata atcaggcaat ctcccaagcc agaatggctc agagagactc 5700
ccaagtccta cattttcaag ttaaaattca ggcagtgaaa ctccagagca cacaactggg 5760
ccatggctgt aaaccagcta cgttctgaat ctcagatcca tgaaaacaca gaggcagcct 5820
caccactccg caggctgtga acaggacgga cggctggcta ctggccatct ccctgcctgc 5880
ctgttgatcc actttcaaca ggcctgagtc tttctctgac tcacccatca ggtggacccc 5940
ctgcagcagg agtgttcccg gccacagaaa gccgattccc tcatctcatc actcaatacc 6000
tatctactga gtcctcctcc tcagagtcca gccctgcgct gggggcacag agaggagtca 6060
gctccccgct gagacagaaa ccaggtactg gctgggcgga gggactgcag gaggggcag 6120
ggtgctgaag aggggtccat ggcagttggt gatggatgag ctgagcctgg gaagggaaca 6180
gaggtgggct agggtctgga gacagacagg gaagcctaga gcagctgaag aacgggaagt 6240
cagtgggaca ggaagacagg aggccgggtt agaaccagat ggccgtcagc cttaactcca 6300
aggaattgaa agccaacatg gttgcatgga tgatgctttg gaagcctggg gggccgggtg 6360
ctagtcttcc ctttctgaag ataacacccc gtcaaggacc gttcacagtg cggacagccc 6420
tggcagcact taaaaggcac cagaagcgaa tgagaagcag aatcgttgat gcgtctgagc 6480
acgtgcacaa agcagacacg gttctcagga ctcgcattcc ttcacccact tcattcccac 6540
aacatcccta agagggaaat acttcaccga tgaggagacc gaggcacaga gcagttatgg 6600
cacgtccccg aggtcacaca gtcagagatg gaacccggac ttgaccccag gcaggctggc 6660
tccagagcat ctccagtgct ccaccacgat gctggaggtc tggggctctt tctcagcagt 6720
tcaactcaga cccagactcc acagcagggc ctgggggcca atcagaatag gccctctcac 6780
ccaaggggct aggcataaga cctgaacctg aagcacctct cgctcaaaca agcagctcag 6840
gaaccccaaa atgcagtgtg aataagaccc aggtcatacc catgcctggc tgtcccaggg 6900
atgtctatgg tctccaagaa taagctaaga ataagaatag gggctgctcg ttaggcagcc 6960
cagcaagaga gcacctgtta gccggagcct cccatgaacc tttcggcaat gttctttgtc 7020
agcggcgttg ccttttcctc tgacgggtcc cagtccaacc aaacccagcc aacaggaagc 7080
agcctggatt gtgctatcat tcagtcagtc gacaaacgct gaatctatac tgcacactct 7140
gcacctgatt ccctactgga aaccaagctt cttccaggga acggggtggg acaggaactg 7200
aggggtcatg tgccctgggc tgggccgggg aacagggtca ggacagccct ccactgccag 7260
gcctggagtg taccctgcac ctaagcccgt ggtgatgcac ctgtcaccgc aaaacattct 7320
acagtgatca ggggaacttc taccgtgatt taggcaaacg agtgttggct gtgtgatgga 7380
gatggtgatt ctggtttaca caagactcac aaactcatca atctggacta aaggacaatc 7440
ctaaaagaat cgattctgcc cacagagacg tggtatgatc tgagtctttc tcaactccaa 7500
agccagaaca gatactgcct accctctggg ggaaggatcc cttctctccc caatgcaagt 7560
gagacaacat ggggtgtggc tctctagaac tagtggatcc cccgggaccg gtcccccac 7620
tcacctgcca gtaagcagtg ggttctctag ttagccagag agctctgctt atatagacct 7680
cccaccgtac acgcctaccg cccatttgcg tcaatggggc ggagttgtta cgacattttg 7740
gaaagtcccg ttgattttgg tgccaaaaca aactcccatt gacgtcaatg gggtggagac 7800
ttggaaatcc ccgtgagtca aaccgctatc cacgcccatt gatgtactgc caaaaccgca 7860
tcaccatggt aatagcgatg actaatacgt agatgtactg ccaagtagga agtcccata 7920
aggtcatgta ctgggcataa tgccaggcgg gccatttacc gtcattgacg tcaatagggg 7980
gcgtacttgg catatgatac acttgatgta ctgccaagtg ggcagtttac cgtaaatagt 8040
ccacccattg acgtcaatgg aaagtcccta ttggcgttac tatgggaaca tacgtcatta 8100
ttgacgtcaa tgggcggggg tcgttgggcg gtcagccagg cgggccattt accgtaagtt 8160
atgtaacgcg gaactccata tatgggctat gaactaatga ccccgtaatt gattactatt 8220
aataactagt caataatcaa tgtcaacgcc cgggctgcag gaattctacc gggtagggga 8280
ggcgcttttc ccaaggcagt ctggagcatg cgctttagca gccccgctgg cacttggcgc 8340
atcacaagtg gcctctggcc tcgcacacat tccacatcca ccggtagcgc caaccggctc 8400
ccttctttgg tggcccttc gcgccacctt ctactcctcc cctagtcagg aagttccccc 8460
ccgccccgca gctcgcgtcg tgcaggacgt gacaaatgga agtagcacgt ctcactagtc 8520
tcgtgcagat ggacaagcac cgctgagcaa tggaagcggg taggcctttg gggcagcggc 8580
caatagcagc ttggctcctt cgctttctgg gctcagaggc tgggaagggg tgggtccggg 8640
ggcgggctca ggggcgggct caggggcggg gcgggcgcga aggtcctccg gacccggcat 8700
tctgcacgct tcaaaagcgc acgtctgccg cgctgttctc ctcttcctca tctccgggcc 8760
ttcgacctgc atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 8820
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcgtgcttt 8880
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 8940
ctacaaagat cgttatgttt atcggcactt tgcatcggcc gcgctcccga ttccggaagt 9000
```

```
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg 9060
tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc cggtcgcgga 9120
ggccatggat g                                                      9131
```

<210> 84
<211> 1121
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      EYPFP sequence

<400> 84
```
cttgtacagc tcgtccatgc cgagagtgat cccggcggcg gtcacgaact ccagcaggac 60
catgtgatcg cgcttctcgt tggggtcttt gctcagggcg gactggtagc tcaggtagtg 120
gttgtcgggc agcagcacgg ggccgtcgcc gatgggggtg ttctgctggt agtggtcggc 180
gagctgcacg ctgccgtcct cgatgttgtg gcggatcttg aagttcacct tgatgccgtt 240
cttctgcttg tcggccatga tatagacgtt gtggctgttg tagttgtact ccagcttgtg 300
ccccaggatg ttgccgtcct ccttgaagtc gatgcccttc agctcgatgc ggttcaccag 360
ggtgtcgccc tcgaacttca cctcggcgcg ggtcttgtag ttgccgtcgt ccttgaagaa 420
gatggtgcgc tcctggacgt agccttcggg catggcggac ttgaagaagt cgtgctgctt 480
catgtggtcg gggtagcggg cgaagcactg caggccgtag ccgaaggtgg tcacgagggt 540
gggccagggc acgggcagct tgccggtggt gcagatgaac ttcagggtca gcttgccgta 600
ggtggcatcg ccctcgccct cgccggacac gctgaacttg tggccgttta cgtcgccgtc 660
cagctcgacc aggatgggca ccaccccggt gaacagctcc tcgcccttgc tcaccatggt 720
ggcgaccggt ggatccttct agagtcgacg atcgaagtta gcttgatttg acagtggctg 780
ggggttgcgc cgcgggttt tataggaagc cacagcggcc actcagcca taaaaggcaa 840
ctttaggaac ggcgggggt gattggattc gagtcgttta ttcaccggcc ttgccgcaca 900
gtgcagcatt tttttacccc ctctcccctc cttttgcggg ggaaaaaaaa aaaaaaaaaa 960
aaaaaaggag aagagaaaaa aagcgagcga gagagaaagc gagattgagg aagaggatga 1020
agagtttggc gatgggtgct ggttccgtag gcccagatgg acaagaatag cccccgccct 1080
tgcggacagt atcccattca gtgactcaga tcagatcaag c                     1121
```

<210> 85
<211> 10557
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      pKI-CMV-SD-Vanilloid-YFP sequence

<400> 85
```
tttttacccc ctctcccctc cttttgcggg ggaaaaaaaa aaaaaaaaaa aaaaaaggag 60
aagagaaaaa aagcgagcga gagagaaagc gagattgagg aagaggatga agagtttggc 120
gatgggtgct ggttccgtag gcccagatgg acaagaatag cccccgccct tgcggacagt 180
atcccattca gtgactcaga tcagatcaag ccggccgctt gccaagggcc ctgtgaagca 240
gggtacccca aaactccaac ttgctgctgc ccccagcccc ctagaactga agtgtttcca 300
ggactgaacc tctcacccca gaggcctcag aggtgagcag gcctagcagg cccaggcag 360
tcctgcggcc actgacgtcg gcctggggagg gtggacaggc ccccggctca gagccatctg 420
tggttctggg ttagacccat ccctcctcgt ctctgggcct caggtgcctc ctctgcaggg 480
tgaggttgat aaccaccgcg ctgcctccct ccccacacag aacaaggaaa acggccgggg 540
tggctgtcgt gggctgaatt cagggaggat ctcggtgatg aggatggctc agtccagaag 600
agccaagcag actgtgtccc tggaataggg tgagcggcag agactggccc acccacaccc 660
acaccgcact gcggggccct gctgtgcaac gggtccacac atcatcagcc aggtgcccgg 720
```

```
cctgcctttg ctcttcaacc tcgctgccac agccactgag cttatacgag attttggatt 780
ttgtgctttt tgtcattttt atgaatcaac atctgtcaag catcctggcc ttgaatgttt 840
aggcgtgtgt gtgtgtgtgt gtgtgtgtgt gtctcactct gtcgcccaga ctagagtgca 900
atggggcaat ctctgttcac tgcaacctcc acctcccagg ttcaagcgat tctcctccct 960
cagcctccca agtagctggg attacaggcg cccgccacca tgcgtggcta atgtttgtat 1020
ttttagtaga gacagggaac gtttaggctt taagtgattt ggagctggtt ggtcttctag 1080
gcctaaacac ctagagatca gctggccgct ggcagcagca tgtgtagttg gacacatgag 1140
ctgatgctat tttccactcc catcctgccc catccccatg cttcccctcc ctctggctgg 1200
gcatccccac tccctggcac ccaggggtgg cccatacatc cctcccatag cctctattcc 1260
catcgggtct ccacctaagg gcttttttcct ctgtctaccc ttgcacaagt cctatcacca 1320
ctcagatgct acttcccctc aggatccagc actgatgccg cctcctccat gaagccccca 1380
gatccctgcc ctgctcatgc tccccgctgt cctccacaca gtggcccagc ccactgaccg 1440
cctgtcccac acagcctacc atgtggtcca gcaattgcta tctgtcactg atttatctct 1500
cccgcaccct gtgacctcct cctgggcagg gacctttcct ggtccatctc cttatgaaga 1560
agctgtactg gcctgggcgc gtggctcacg cctgtaatcc cagcactttg ggaggccaag 1620
gcaggtggat cacctgaggt caggagttcg agaccagcct ggccaacatg gtgaaatccc 1680
atctctacta aaaaaaatac aaaaattagc cgggcgtggt gtcgggcgcc tgtaatccca 1740
gctactcggg aggctgaggc aggagaatcg cttgaaccca ggatgtagag gttacagtga 1800
gctgagactg tgccactgca ctccagcctg ggcaacagag tgaaactctg tctcaaaaaa 1860
aaaaaaaaaa aaaaaaaaag aagcagcagc agcagcagca gcagcagctg cacacatggg 1920
gaaataagag caatcagacc cttctaattg ccattccat aggaagttag ggagattact 1980
gttccagcag tggctacatt ccggaaaaca gccgcagaag aaatgagagg cttctcaagt 2040
cctgcatttt cttttttttt tttttgaaa tatataatca ggcaatctcc caagccagaa 2100
tggctcagag agactcccaa gtcctacatt ttcaagttaa aattcaggca gtgaaactcc 2160
agagcacaca actgggccat ggctgtaaac cagctacgtt ctgaatctca gatccatgaa 2220
aacacagagg cagcctcacc actccgcagg ctgtgaacag gacggacggc tggctactgg 2280
ccatctccct gcctgcctgt tgatccactt tcaacaggcc tgagtctttc tctgactcac 2340
ccatcaggtg gaccccctgc agcaggagtg ttcccggcca cagaaagccg attccctcat 2400
ctcatcactc aatacctatc tactggctcc tcctcctcag agtccagccc tgcgctgggg 2460
gcacagagag gagtcagctc cccgctgaga cagaaaccag gtactggctg ggcggaggga 2520
ctgcaggagg gggcagggtg ctgaagaggg gtccatggca gttggtgatg gatgagctga 2580
gcctgggaag ggaacagagg tgggctaggg tctggagaca gacagggaag cctagagcag 2640
ctgaagaacg ggaagtcagt gggacaggaa gacaggaggc cgggttagaa ccagatggcc 2700
gtcagcctta actccaagga attgaaagcc aacatggttg catggatgat gctttggaag 2760
cctggggggc cgggtgctag tcttcccttt ctgaagataa caccccgtca aggaccgttc 2820
acagtgcgga cagccctggc agcacttaaa aggcaccaga agcgaatgag aagcagaatc 2880
gttgatgcgt ctgagcacgt gcacaaagca gacacggttc tcaggactcg cattccttca 2940
cccacttcat tcccacaaca tccctaagag ggaaatactt caccgatgag gagaccgagg 3000
cacagagcag ttatggcacg tccccgaggt cacacagtca gagatggaac ccggacttga 3060
ccccaggcag gctggctcca gagcatctcc agtgctccac cacgatgctg gaggtctggg 3120
gctctttctc agcagttcaa ctcagaccca gactccacag cagggcctgg gggccaatca 3180
gaataggccc tctcacccaa ggggctaggc ataagacctg aacctgaagc acctctcgct 3240
caaacaagca gctcaggaac cccaaaatgc agtgtgaata agacccaggt catacccatg 3300
cctggctgtc ccagggatgt ctatggtctc caagaataag ctaagaataa gaataggggc 3360
tgctcgttag gcagcccagc aagagagcac ctgttagccg gagcctccca tgaacctttc 3420
ggcaatgttc tttgtcagcg gcgttgcctt ttcctctgac gggtcccagt ccaaccaaac 3480
ccagccaaca ggaagcagcc tggattgtgc tatcattcag tcagtcgaca aacgctgaat 3540
ctatactgca cactctgcac ctgattccct actggaaacc aagcttcttc cagggaacgg 3600
ggtgggacag gaactgaggg gtcatgtgcc ctgggctggg ccggggaaca gggtcaggac 3660
agccctccac tgccaggcct ggagtgtacc ctgcacctaa gcccgtggtg atgcacctgt 3720
caccgcaaaa cattctacag tgatcagggg aacttctacc gtgatttagg caaacgagtg 3780
ttggctgtgt gatggagatg gtgattctgg tttacacaag actcacaaac tcatcaatct 3840
ggactaaagg acaatcctaa aagaatcgat tctgcccaca gagacgtggt atgatctgag 3900
tctttctcaa ctccaaagcc agaacagata ctgcctaccc tctgggggaa ggatcccttc 3960
tctccccaat gcaagtgaga caacatgggg tgtggctctc tagaactagt ggatcccccg 4020
ggaccggtcc ccccactcac ctgccagtaa gcagtgggtt ctctagttag ccagagagct 4080
ctgcttatat agacctccca ccgtacacgc ctaccgccca tttgcgtcaa tggggcggag 4140
ttgttacgac attttggaaa gtcccgttga ttttggtgcc aaaacaaact cccattgacg 4200
```

```
tcaatggggt ggagacttgg aaatccccgt gagtcaaacc gctatccacg cccattgatg 4260
tactgccaaa accgcatcac catggtaata gcgatgacta atacgtagat gtactgccaa 4320
gtaggaaagt cccataaggt catgtactgg gcataatgcc aggcgggcca tttaccgtca 4380
ttgacgtcaa tagggggcgt acttggcata tgatacactt gatgtactgc caagtgggca 4440
gtttaccgta aatagtccac ccattgacgt caatggaaag tccctattgg cgttactatg 4500
ggaacatacg tcattattga cgtcaatggg cggggggtcgt tgggcggtca gccaggcggg 4560
ccatttaccg taagttatgt aacgcggaac tccatatatg ggctatgaac taatgacccc 4620
gtaattgatt actattaata actagtcaat aatcaatgtc aacgcccggg ctgcaggaat 4680
tctaccgggt aggggaggcg cttttcccaa ggcagtctgg agcatgcgct ttagcagccc 4740
cgctggcact tggcgcatca caagtggcct ctggcctcgc acacattcca catccaccgg 4800
tagcgccaac cggctccctt ctttggtggc cccttcgcgc caccttctac tcctcccta 4860
gtcaggaagt tcccccccgc cccgcagctc gcgtcgtgca ggacgtgaca aatggaagta 4920
gcacgtctca ctagtctcgt gcagatggac aagcaccgct gagcaatgga agcgggtagg 4980
cctttggggc agcggccaat agcagcttgg ctccttcgct ttctgggctc agaggctggg 5040
aaggggtggg tccggggggcg ggctcagggg cgggctcagg ggcggggcgg gcgcgaaggt 5100
cctccggacc cggcattctg cacgcttcaa aagcgcacgt ctgccgcgct gttctcctct 5160
tcctcatctc cgggccttcg acctgcatga aaaagcctga actcaccgcg acgtctgtcg 5220
agaagtttct gatcgaaaag ttcgacagcg tctccgacct gatgcagctc tcggagggcg 5280
aagaatctcg tgctttcagc ttcgatgtag gagggcgtgg atatgtcctg cgggtaaata 5340
gctgcgccga tggtttctac aaagatcgtt atgtttatcg gcactttgca tcggccgcgc 5400
tcccgattcc ggaagtgctt gacattgggg aattcagcga gagcctgacc tattgcatct 5460
cccgccgtgc acagggtgtc acgttgcaag acctgcctga aaccgaactg cccgctgttc 5520
tgcagccggt cgcggaggcc atggatgcga tcgctgcggc cgatcttagc cagacgagcg 5580
ggttcggccc attcggaccg caaggaatcg gtcaatacac tacatggcgt gatttcatat 5640
gcgcgattgc tgatccccat gtgtatcact ggcaaactgt gatggacgac accgtcagtg 5700
cgtccgtcgc gcaggctctc gatgagctga tgctttgggc cgaggactgc cccgaagtcc 5760
ggcacctcgt gcacgcggat ttcggctcca acaatgtcct gacggacaat ggccgcataa 5820
cagcggtcat tgactggagc gaggcgatgt tcggggattc ccaatacgag gtcgccaaca 5880
tcttcttctg gaggccgtgg ttggcttgta tggagcagca gacgcgctac ttcgagcgga 5940
ggcatccgga gcttgcagga tcgccgcggc tccgggcgta tatgctccgc attggtcttg 6000
accaactcta tcagagcttg gttgacggca atttcgatga tgcagcttgg gcgcagggtc 6060
gatgcgacgc aatcgtccga tccggagccg ggactgtcgg gcgtacacaa atcgcccgca 6120
gaagcgcggc cgtctggacc gatggctgtg tagaagtact cgccgatagt ggaaaccgac 6180
gccccagcac tcgtccgagg gcaaaggaat gagctcgctg atcagcctcg actgtgcctt 6240
ctagttgcca gccatctgtt gtttgcccct cccccgtgcc ttccattgt ctgagtaggt 6300
gtcattctat tctgggggggt ggggtggggc aggacagcaa ggggggaggat tgggaagaca 6360
atagcaggca tgctggggat gcggtgggct ctatggcttc tgaggcggaa agaaccagct 6420
ggggggatatc aagcttatcg ataccgtcga cctcgagggg gggcccggta cccagctttt 6480
gttcccttta gtgagggtta attgcgcgct tggcgtaatc atggtcatag ctgtttcctg 6540
tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgta 6600
aagcctgggg tgcctaatga gtgagctaac tcacattaat tgcgttgcgc tcactgcccg 6660
ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga 6720
gaggcggttt gcgtattggg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg 6780
tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag 6840
aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc 6900
gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg ccccctgac gagcatcaca 6960
aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt 7020
ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 7080
tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc 7140
tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc 7200
ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact 7260
tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg 7320
ctacagagtt cttgaagtgg tggcctaact acggctacac tagaaggaca gtatttggta 7380
tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca 7440
aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa 7500
aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg 7560
aaaactcacg ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc 7620
ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg 7680
```

```
acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat 7740
ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg 7800
gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa 7860
taaaccagcc agccggaagg gccgagcgca gaagtggtcc tgcaacttta tccgcctcca 7920
tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc 7980
gcaacgttgt tgccattgct acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 8040
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 8100
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 8160
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 8220
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 8280
gttgctcttg cccggcgtca atacgggata ataccgcgcc acatagcaga actttaaaag 8340
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 8400
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 8460
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 8520
cgacacggaa atgttgaata ctcatactct tcctttttca atattattga agcatttatc 8580
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 8640
gggttccgcg cacatttccc cgaaaagtgc cacctaaatt gtaagcgtta atatttttgtt 8700
aaaattcgcg ttaaattttt gttaaatcag ctcattttttt aaccaatagg ccgaaatcgg 8760
caaaatccct tataaatcaa aagaatagac cgagataggg ttgagtgttg ttccagtttg 8820
gaacaagagt ccactattaa agaacgtgga ctccaacgtc aaagggcgaa aaaccgtcta 8880
tcagggcgat ggcccactac gtgaaccatc accctaatca agttttttgg ggtcgaggtg 8940
ccgtaaagca ctaaatcgga accctaaagg gagcccccga tttagagctt gacggggaaa 9000
gccggcgaac gtggcgagaa aggaagggaa gaaagcgaaa ggagcgggcg ctaggcgct 9060
ggcaagtgta gcggtcacgc tgcgcgtaac caccacaccc gccgcgctta atgcgccgct 9120
acagggcgcg tcccattcgc cattcaggct gcgcaactgt tgggaagggc gatcggtgcg 9180
ggcctcttcg ctattacgcc agctggcgaa aggggatgt gctgcaaggc gattaagttg 9240
ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg agcgcgcgta 9300
atacgactca ctatagggcg aattggagct ccaccgcggt ggcggccgca acagacatga 9360
taagatacat tgatgagttt ggacaaacca caactagaat gcagtgaaaa aaatgcttta 9420
tttgtgaaat ttgtgatgct attgctttat ttgtaaccat tataagctgc aataaacaag 9480
ttggggtggg cgaagaactc cagcatgaga tccccgcgct ggaggatcat ccagccgtcg 9540
tcccggaaaa cgattccgaa gcccaacctt tcatagaagg cggcggtgga atcgaaatct 9600
cgtgatggca ggttgggcgt cgcttggtcg gtcatttcga agctttactt gtacagctcg 9660
tccatgccga gagtgatccc ggcggcggtc acgaactcca gcaggaccat gtgatcgcgc 9720
ttctcgttgg ggtctttgct cagggcggac tggtagctca ggtagtggtt gtcgggcagc 9780
agcacggggc cgtcgccgat ggggggtgttc tgctggtagt ggtcggcgag ctgcacgctg 9840
ccgtcctcga tgttgtggcg gatcttgaag ttcaccttga tgccgttctt ctgcttgtcg 9900
gccatgatat agacgttgtg gctgttgtag ttgtactcca gcttgtgccc caggatgttg 9960
ccgtcctcct tgaagtcgat gcccttcagc tcgatgcggt tcaccagggt gtcgccctcg 10020
aacttcacct cggcgcgggt cttgtagttg ccgtcgtcct tgaagaagat ggtgcgctcc 10080
tggacgtagc cttcgggcat ggcggacttg aagaagtcgt gctgcttcat gtggtcgggg 10140
tagcgggcga agcactgcag gccgtagccg aaggtggtca cgagggtggg ccagggcacg 10200
ggcagcttgc cggtggtgca gatgaacttc agggtcagct tgccgtaggt ggcatcgccc 10260
tcgccctcgc cggacacgct gaacttgtgg ccgtttacgt cgccgtccag ctcgaccagg 10320
atgggcacca ccccggtgaa cagctcctcg cccttgctca ccatggtggc gaccggtgga 10380
tccttctaga gtcgacgatc gaagttagct tgatttgaca gtggctgggg gttgcgccgc 10440
cgggttttat aggaagccac agcggccact cgagccataa aaggcaactt taggaacggc 10500
gggggtgat tggattcgag tcgtttattc accggccttg ccgcacagtg cagcatt 10557
```

<210> 86
<211> 3799
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic nucleotide
      sequence for Vanilloid region of homology

<400> 86
```
ttcaagcgat tctcctgcct cagcctcccg agtagctggg attacaggcg cccgacacca  60
cgcccggcta attttttgtat ttttttttagt agagatggga tttcaccatg ttggccaggc 120
tggtctcgaa ctcctgacct caggtgatcc acctgccttg gcctcccaaa gtgctgggat 180
tacaggcgtg agccacgcgc ccaggccagt acagcttctt cataaggaga tggaccagga 240
aaggtccctg cccaggagga ggtcacaggg tgcgggagag ataaatcagt gacagatagc 300
aattgctgga ccacatggta ggctgtgtgg gacaggcggt cagtgggctg ggccactgtg 360
tggaggacag cggggagcat gagcagggca gggatctggg ggcttcatgg aggaggcggc 420
atcagtgctg gatcctgagg ggaagtagca tctgagtggt gataggactt gtgcaagggt 480
agacagagga aaaagccctt aggtggagac ccgatgggaa tagaggctat gggagggatg 540
tatgggccac ccctgggtgc cagggagtgg ggatgcccag ccagagggag gggaagcatg 600
gggatggggc aggatgggag tggaaaatag catcagctca tgtgtccaac tacacatgct 660
gctgccagcg gccagctgat ctctaggtgt ttaggcctag aagaccaacc agctccaaat 720
cacttaaagc ctaaacgttc cctgtctcta ctaaaaatac aaacattagc cacgcatggt 780
ggcgggcgcc tgtaatccca gctacttggg aggctgaggg aggagaatcg cttgaacctg 840
ggaggtggag gttgcagtga acagagattg ccccattgca ctctagtctg ggcgacagag 900
tgagacacac acacacacac acacacac acgcctaaac attcaaggcc aggatgcttg 960
acagatgttg attcataaaa atgacaaaaa gcacaaaatc caaaatctcg tataagctca 1020
gtggctgtgg cagcgaggtt gaagagcaaa ggcaggcgg gcacctggct gatgatgtgt 1080
ggacccgttg cacagcaggg ccccgcagtg cggtgtgggt gtgggtgggc cagtctctgc 1140
cgctcaccct attccaggga cacagtctgc ttggctcttc tggactgagc catcctcatc 1200
accgagatcc tccctgaatt cagcccacga cagccacccc ggccgttttc cttgttctgt 1260
gtggggaggg aggcagcgcg gtggttatca acctcaccct gcagaggagg cacctgaggc 1320
ccagagacga ggagggatgg gtctaaccca gaaccacaga tggctctgag ccggggggcct 1380
gtccaccctc ccaggccgac gtcagtggcc gcaggactgc ctgggccctg ctaggcctgc 1440
tcacctctga ggcctctggg gtgagaggtt cagtcctgga aacacttcag ttctaggggg 1500
ctgggggcag cagcaagttg gagttttggg gtaccctgct tcacagggcc cttggcaagc 1560
ggccgctcag atctagagag ccacacccca tgttgtctca cttgcattgg ggagagaagg 1620
gatccttccc ccagagggta ggcagtatct gttctggctt tggagttgag aaagactcag 1680
atcataccac gtctctgtgg gcagaatcga ttctttaggg attgtccttt agtccagatt 1740
gatgagtttg tgagtcttgt gtaaaccaga atcaccatct ccatcacaca gccaacactc 1800
gtttgcctaa atcacggtag aagttcccct gatcactgta gaatgttttg cggtgacagg 1860
tgcatcacca cgggcttagg tgcagggtac actccaggcc tggcagtgga gggctgtcct 1920
gaccctgttc cccggcccag cccagggcac atgacccctc agttcctgtc ccaccccgtt 1980
ccctggaaga agcttggttt ccagtaggga atcaggtgca gagtgtgcag tatagattca 2040
gcgtttgtcg actgactgaa tgatagcaca atccaggctg cttcctgttg gctgggtttg 2100
gttggactgg gacccgtcag aggaaaaggc aacgccgctg acaaagaaca ttgccgaaag 2160
gttcatggga ggctccggct aacaggtgct ctcttgctgg gctgcctaac gagcagcccc 2220
tattcttatt cttagcttat tcttggagac catagacatc cctgggacag ccaggcatgg 2280
gtatgacctg ggtcttattc acactgcatt ttggggttcc tgagctgctt gtttgagcga 2340
gaggtgcttc aggttcaggt cttatgccta gcccccttggg tgagagggcc tattctgatt 2400
ggcccccagg ccctgctgtg gagtctgggt ctgagttgaa ctgctgagaa agagccccag 2460
acctccagca tcgtggtgga gcactggaga tgctctggag ccagcctgcc tggggtcaag 2520
tccgggttcc atctctgact gtgtgacctc ggggacgtgc cataactgct ctgtgcctcg 2580
gtctcctcat cggtgaagta tttccctctt agggatgttg tgggaatgaa gtgggtgaag 2640
gaatgcgagt cctgagaacc gtgtctgctt tgtgcacgtg ctcagacgca tcaacgattc 2700
tgcttctcat tcgcttctgg tgcctttttaa gtgctgccag ggctgtccgc actgtgaacg 2760
gtccttgacg gggtgttatc ttcagaaagg gaagactagc acccggcccc ccaggcttcc 2820
aaagcatcat ccatgcaacc atgttggctt tcaattcctt ggagttaagg ctgacggcca 2880
tctggttcta acccggcctc ctgtcttcct gtcccactga cttcccgttc ttcagctgct 2940
ctaggcttcc ctgtctgtct ccagaccctta gcccacctct gttcccttcc caggctcagc 3000
tcatccatca ccaactgcca tggaccccctc ttcagcaccc tgcccctcc tgcagtccct 3060
ccgcccagcc agtacctggt ttctgtctca gcggggagct gactcctctc tgtgccccca 3120
gcgcaggggct ggactctgag gaggaggact cagtagatag gtattgagtg atgagatgag 3180
ggaatcggct ttctgtggcc gggaacactc ctgctgcagg gggtccacct gatgggtgag 3240
tcagagaaag actcaggcct gttgaaagtg gatcaacagg caggcaggga gatggccagt 3300
agccagccgt ccgtcctgtt cacagcctgc ggagtggtga ggctgcctct gtgttttcat 3360
```

```
ggatctgaga ttcagaacgt agctggttta cagccatggc ccagttgtgt gctctggagt 3420
ttcactgcct gaattttaac ttgaaaatgt aggacttggg agtctctctg agccattctg 3480
gcttgggaga ttgcctgatt atatatttca aaaaaaaaaa aaaagaaaat gcaggacttg 3540
agaagcctct catttcttct gcggctgttt tccggaatgt agccactgct ggaacagtaa 3600
tctccctaac ttcctatgga aatggcaatt agaagggtct gattgctctt atttccccat 3660
gtgtgcagct gctgctgctg ctgctgctgc tgcttctttt tttttttttt tttttttttt 3720
tgagacagag tttcactctg ttgcccaggc tggagtgcag tggcacagtc tcagctcact 3780
gtaacctcta catcctggg                                            3799
```

US 107084200QP1

Creation date: 09-08-2004
Indexing Officer: SGIRON - SERGIO GIRON
Team: OIPEBackFileIndexing
Dossier: 10708420

Legal Date: 07-15-2004

No.	Doccode	Number of pages
1	PEFR	5
2	OATH	2
3	A.PE	1
4	CLM	5
5	REM	1
6	IDS	4
7	FOR	68
8	FOR	11
9	FOR	22
10	FOR	16
11	NPL	12
12	NPL	12

Total number of pages: 159

Remarks:

Order of re-scan issued on ....